

STIC-Biotech/ChemLib

85948

From: Borin, Michael
Sent: Tuesday, February 04, 2003 11:14 AM
To: STIC-Biotech/ChemLib
Subject: RE: Search request: 09/865548

Examiner: M.Borin
CM1 12A01
AU: 1631; Mailbox 12D01

Tel.: 305-4506

RE: 09/865548; peptide

Please conduct search of polypeptide SEQ ID 13 against the commercial and interference protein databases.

Thank you

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

Searcher: Larson
Phone: _____
Location: _____
Date Picked Up: 2/4
Date Completed: 2/4
Searcher Prep/Review: 5
Clerical: _____
Online time: 5

TYPE OF SEARCH:
NA Sequences: _____
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: ABSS 02
WWW/Internet: _____
Other (specify): _____



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:36:35 ; Search time 35 Seconds
(without alignments)
34.264 Million cell updates/sec

Title: US-09-865-548A-13
Perfect score: 43
Sequence: 1 GLIEKNTIEL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	43	100.0	1042	22	ABG14395
2	43	100.0	1080	22	Novel human diagno
3	43	100.0	1603	23	Novel human diagno
4	37	86.0	139	23	Human polypeptide
5	35	81.4	269	22	Streptococcus poly
6	34	79.1	254	22	Drosophila melanog
7	34	79.1	360	22	Enterococcus faeca
8	34	79.1	974	23	Pentaria chalcone s
9	33	76.7	175	21	Fungal ZBC protein
10	33	76.7	203	21	Arabidopsis thalia

11	33	76.7	306	21	AAG17120
12	33	76.7	974	19	AAW55960
13	33	76.7	1056	22	ABB69318
14	32	74.4	122	23	ABP38169
15	32	74.4	235	22	AAU20516
16	32	74.4	352	21	AAAG14893
17	32	74.4	352	21	AAAG48119
18	32	74.4	610	21	AAAG18247
19	32	74.4	1381	22	ABBS7920
20	31	72.1	91	22	ABG11365
21	31	72.1	91	22	ABG12566
22	31	72.1	106	21	AAAG01367
23	31	72.1	111	23	ABP32374
24	31	72.1	136	22	AAUG7663
25	31	72.1	185	21	AAAB42077
26	31	72.1	216	23	ABBS5277
27	31	72.1	232	20	AAU77257
28	31	72.1	235	23	ABP26523
29	31	72.1	237	23	ABP41753
30	31	72.1	238	22	ABG19263
31	31	72.1	280	10	AAAP95648
32	31	72.1	397	21	AAAY45082
33	31	72.1	428	23	ABP38097
34	31	72.1	859	21	AAAY96997
35	31	72.1	1037	22	ABBS3585
36	30	69.8	53	21	AAAY50724
37	30	69.8	57	22	ABBI4811
38	30	69.8	58	21	AAAG24970
39	30	69.8	110	23	ABP28503
40	30	69.8	149	23	ABBS4299
41	30	69.8	164	23	AAAM49533
42	30	69.8	166	23	ABP34827
43	30	69.8	167	23	ABP42602
44	30	69.8	201	21	AAAG24096
45	30	69.8	213	21	AAAG39139

ALIGNMENTS

RESULT 1
ABG14395
ID ABG14395 standard; Protein; 1042 AA.
XX AC
XX ABG14395;
XX AC
DT 18-FEB-2002 (first entry)
XX
XX Novel human diagnostic protein #14386.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.

XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US08631.
XX 31-MAR-2000; 2000US-0540217.
XX 23-AUG-2000; 2000US-0649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS/8582.

XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations

Arabidopsis thalia
Human transient re
Drosophila melanog
Staphylococcus epi
Human secreted pro
Arabidopsis thalia
Arabidopsis thalia
Plasmodium falcipa
Drosophila melanog
Novel human diagno
Novel human diagno
Human secreted pro
Human ORF1347 prot
Propionibacterium
Human ORF ORF1841
Lactococcus lactis
Human adenylate ki
Streptococcus poly
Human ovarian anti
Novel human diagno
Ricin agglutinin A
Rice sulphite oxid
Staphylococcus
S. cerevisiae e
Human protein seq
S. aureus sarA pro
Human nervous syst
Arabidopsis thalia
Streptococcus poly
Lactococcus lactis
S. synthia telomer
Human polymerase-1
Human ovarian anti
Arabidopsis thalia
Arabidopsis thalia

PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 44754; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Query Match 100.0%; Score 43; DB 22; Length 1042;
SQ Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
Db 345 GLIEKNIEL 353
|||||

RESULT 2
ABG14396
ID ABG14396 standard; Protein; 1080 AA.
AC
XX ABG14396;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #14387.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US08631.
XX
XX 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS78583.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

XX Claim 20; SEQ ID No 44755; 103pp; English.

PS The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Query Match 100.0%; Score 43; DB 22; Length 1080;
SQ Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
Db 315 GLIEKNIEL 323
|||||

RESULT 3
AAM40302
ID AAM40302 standard; Protein; 1603 AA.
XX
XX AAM40302;
AC
XX 22-OCT-2001 (first entry)
DT
XX
DE Human polypeptide SEQ ID NO 3447.
XX
XX
KW Human; neurotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US34263.
PF
XX
XX 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-052317.
PR 09-JUL-2000; 2000US-0598042.
PR 19-JUL-2000; 2000US-0620312.
PR 03-AUG-2000; 2000US-0653450.
PR 14-SEP-2000; 2000US-0662191.
PR 19-OCT-2000; 2000US-0693036.
PR 29-NOV-2000; 2000US-0727344.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI

PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR N-PSDB; AAI59458.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 6; SEQ ID NO 3447; 10078pp; English.
 XX
 CC The invention relates to human nucleic acids (AAI57798-AAI61369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 1603 AA;
 Query Match 100.0%; Score 43; DB 22; Length 1603;
 Best Local Similarity 100.0%; Pred. NO. 9.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GLIEKNIEL 9
 Db 425 GLIEKNIEL 433
 |||||
 RESULT 4
 ABP26289
 ID ABP26289 standard; Protein; 139 AA.
 AC ABP26289;
 XX
 DT 02-JUL-2002 (first entry)
 DE Streptococcus polypeptide SEQ ID NO 1754.
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 FI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tettelin H;
 XX
 DR WPI: 2002-352536/38.
 DR N-PSDB; ABN66920.

XX New Streptococcus protein for the treatment or prevention of infection
 PT or disease caused by Streptococcus bacteria, such as meningitis, and
 PT for detecting a compound that binds to the protein -
 XX
 PS Claim 1; Page 3325; 4525pp; English.
 XX
 CC The invention relates to a protein (ABP25413-ABP30895) from group B
 CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
 CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
 CC the specification. The proteins have antibacterial and antiinflammatory
 CC activity. (I), nucleic acids encoding (I), ABN66044-ABN71526 and
 CC antibodies that bind (I) are used in the manufacture of medicaments for
 CC the treatment or prevention of infection or disease caused by
 CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
 CC Nucleic acids encoding (I) are used to detect Streptococcus in a
 CC biological sample. (I) is used to determine whether a compound binds to
 CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
 CC used as a vaccine or diagnostic composition. The disease caused by
 CC Streptococcus that is prevented or treated may be meningitis. Nucleic
 CC acid encoding (I) may be used to recombinantly produce (I) and may be
 CC used in gene therapy. Antibodies to (I) are used for affinity
 CC chromatography, immunoassays, and distinguishing/identifying
 CC Streptococcus proteins.
 XX
 SQ Sequence 139 AA;
 Query Match 86.0%; Score 37; DB 23; Length 139;
 Best Local Similarity 88.9%; Pred. NO. 11;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GLIEKNIEL 9
 Db 100 GLIEKNIEL 108
 |||||
 RESULT 5
 ABB67829
 ID ABB67829 standard; Protein; 269 AA.
 AC ABB67829;
 XX
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 30279.
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR N-PSDB; ABL11932.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 30279; 21pp + Sequence Listing; English.
 XX

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB557737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 269 AA;

Query Match 81.4%; Score 35; DB 22; Length 269;
 Best Local Similarity 87.5%; Pred. No. 54;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
 |||||
 Db 23 LIEKNIEI 30

RESULT 6

AAU35137
 ID AAU35137 standard; Protein; 254 AA.

XX
 AC AAU35137;

DT 13-FEB-2002 (first entry)

DE Enterococcus faecalis cellular proliferation protein #424.

XX Antisense; prokaryotic cellular proliferation protein;
 KW antibiotic; antibacterial; drug design.

OS Enterococcus faecalis.

XX WO200170955-A2.

PN 27-SEP-2001.

XX 21-MAR-2001; 2001WO-US09180.

PF 21-MAR-2000; 2000US-191078P.

PR 23-MAY-2000; 2000US-206848P.

PR 26-MAY-2000; 2000US-207727P.

PR 23-OCT-2000; 2000US-242578P.

PR 27-NOV-2000; 2000US-253625P.

PR 22-DEC-2000; 2000US-257931P.

PR 16-FEB-2001; 2001US-269308P.

XX (ELIT-) ELITRA PHARM INC.

XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;

PI Yamamoto RT, Xu HH;

XX WPI: 2001-611495/70.

DR N-PSDB; AAS52996.

XX New polynucleotides for the identification and development of

PT antibiotics, comprise sequences of antisense nucleic acids -

XX Example 3; Seq ID No 10730; 511pp; English.

XX The invention relates to antisense inhibitors of genes essential to

CC prokaryotic cellular proliferation, their use in identifying the

CC genes, their use in the discovery of novel antibiotics, the essential

CC genes themselves and the encoded proteins. The prokaryotes used are

CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 254 AA;

Query Match 79.1%; Score 34; DB 22; Length 254;
 Best Local Similarity 75.0%; Pred. No. 81;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps

QY 2 LIEKNIEL 9
 |||||
 Db 167 LIEKNVEI 174

RESULT 7

AAAB60169
 ID AAAB60169 standard; Protein; 360 AA.

XX
 AC AAAB60169;

DT 03-APR-2001 (first entry)

DE Petunia chalcone synthase.

XX Petunia; chalcone synthase; reversible male sterility; yield; uniformity;
 KW Ms*5126; self-pollination.

OS Petunia hybrida.

XX WO200100834-A1.

PN 04-JAN-2001.

XX 27-JUN-2000; 2000WO-US17624.

XX 29-JUN-1999; 99US-0340684.

XX (PION-) PIONEER HI-BRED INT INC.

XX Fox T, Albertsen MC;

XX WPI: 2001-112456/12.

XX Novel plants useful in plant hybridization techniques comprise an

PT endogenous Ms.5126 gene, the expression of which is impaired such that

PT the plant possesses male sterility -

XX Disclosure; Fig 2; 53pp; English.

XX The present invention describes a plant containing an endogenous Ms*5126

CC gene, the expression of which is impaired, causing male fertility in the

CC plant to be impaired. This is useful in the production of male sterile

CC plants. These are useful in the production of hybrid seed, which leads to

CC plants superior in vigour, yield and uniformity, making them more

CC attractive to farmers.

XX Sequence 360 AA;

Query Match 79.1%; Score 34; DB 22; Length 360;

Best Local Similarity 87.5%; Pred. No. 1.2e+02;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
 |||||

```
Db 289 GLISKNE 296
RESULT 8
ABP35581
ID ABP35581 standard; Protein; 974 AA.
XX
AC ABP35581;
XX
DT 24-JUL-2002 (first entry)
XX
DE Fungal ZBC protein sequence #7.
XX
KW Secondary metabolite; fungus; ZBC gene; zinc binuclear cluster protein;
KW antibacterial; beta-lactam; anti-hypercholesterolaemic; lovastatin;
KW mevastatin; immunosuppressant; cyclosporin A; ergot alkaloid; ergotamine;
KW angioestasis inhibitor; ovalicin; glucan synthase inhibitor; gliotoxin;
KW fungal toxin; cell surface receptor; plant growth regulator; pigment;
KW insecticide; antineoplastic.
XX
OS Unidentified.
XX
PN WO200224865-A2.
XX
PD 28-MAR-2002.
XX
PF 19-SEP-2001; 2001WO-US29288.
XX
PR 19-SEP-2000; 2000US-233564P.
XX
PA (MICR-) MICROBIA INC.
XX
PI Holtzman D, Madden K, Maxon M, Sherman A;
XX
DR WP1: 2002-352005/38.
XX
N-PSDB; ABN79770.
XX
PT New method for improving the production of a secondary metabolite e.g.
PT antineoplastic agent, ergot alkaloid from a fungus involves modulation
PT of the expression of at least one zinc binuclear cluster protein gene
PT .
XX
PS Disclosure; SEQ ID 26; 49pp + sequence listing; English.
CC
CC The invention relates to improving the production of a secondary
CC metabolite by a fungus. This involves modulating the expression of at
CC least one ZBC (zinc binuclear cluster protein) gene in a manner to
CC improve the yield of the secondary metabolite. Methods of the invention
CC may be used for improving the production of the secondary metabolite e.g.
CC antibacterial (such as beta-lactam), an anti-hypercholesterolaemic (such
CC as lovastatin or mevastatin), an immunosuppressant (such as cyclosporin A),
CC an ergot alkaloid (such as ergotamine), an angioestasis inhibitor (such
CC as ovalicin), a glucan synthase inhibitor, gliotoxin family of compounds,
CC a fungal toxin, a modulator of cell surface receptor signalling, a plant
CC growth regulator, a pigment, an insecticide, or an antineoplastic
CC compound. The method results in a decrease in fermentor run-time, a
CC decrease in the size of the fermentor required for the production of
CC equivalent amounts of the secondary metabolite, or a decrease in the
CC biomass required for the production, which translates into decreased
CC waste that must be handled in downstream processing. The sequences given
CC in records ABP35575-ABP35722 represent ZBC proteins.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 974 AA;
Query Match 79.1%; Score 34; DB 23; Length 974;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 2 LIEKNIEL 9
:|||||
```

```
Db 519 IIEKNVEL 526
RESULT 9
AAG17122
ID AAG17122 standard; Protein; 175 AA.
XX
AC AAG17122;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18024.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
```

PR 18-JUN-1999;	99US-0139456.	PR 18-AUG-1999;	99US-0149426.
PR 18-JUN-1999;	99US-0139457.	PR 20-AUG-1999;	99US-0149722.
PR 18-JUN-1999;	99US-0139458.	PR 20-AUG-1999;	99US-0149723.
PR 18-JUN-1999;	99US-0139459.	PR 20-AUG-1999;	99US-0149929.
PR 18-JUN-1999;	99US-0139460.	PR 23-AUG-1999;	99US-0149902.
PR 18-JUN-1999;	99US-0139461.	PR 23-AUG-1999;	99US-0149930.
PR 18-JUN-1999;	99US-0139462.	PR 25-AUG-1999;	99US-0150566.
PR 18-JUN-1999;	99US-0139463.	PR 26-AUG-1999;	99US-0150884.
PR 18-JUN-1999;	99US-0139750.	PR 27-AUG-1999;	99US-0151065.
PR 18-JUN-1999;	99US-0139763.	PR 27-AUG-1999;	99US-0151066.
PR 21-JUN-1999;	99US-0139817.	PR 27-AUG-1999;	99US-0151080.
PR 22-JUN-1999;	99US-0139899.	PR 30-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140353.	PR 31-AUG-1999;	99US-0151303.
PR 23-JUN-1999;	99US-0140354.	PR 07-SEP-1999;	99US-0151930.
PR 24-JUN-1999;	99US-0140695.	PR 10-SEP-1999;	99US-0153070.
PR 28-JUN-1999;	99US-0140823.	PR 13-SEP-1999;	99US-0153758.
PR 29-JUN-1999;	99US-0140991.	PR 15-SEP-1999;	99US-0154018.
PR 30-JUN-1999;	99US-0141287.	PR 16-SEP-1999;	99US-0154039.
PR 01-JUL-1999;	99US-0141842.	PR 20-SEP-1999;	99US-0154779.
PR 01-JUL-1999;	99US-0142154.	PR 22-SEP-1999;	99US-0155139.
PR 02-JUL-1999;	99US-0142055.	PR 23-SEP-1999;	99US-0155486.
PR 06-JUL-1999;	99US-0142390.	PR 24-SEP-1999;	99US-0155659.
PR 08-JUL-1999;	99US-0142803.	PR 28-SEP-1999;	99US-0156458.
PR 09-JUL-1999;	99US-0142920.	PR 04-OCT-1999;	99US-0157117.
PR 12-JUL-1999;	99US-0142977.	PR 05-OCT-1999;	99US-0157753.
PR 13-JUL-1999;	99US-0143542.	PR 06-OCT-1999;	99US-0157865.
PR 14-JUL-1999;	99US-0143624.	PR 07-OCT-1999;	99US-0158029.
PR 15-JUL-1999;	99US-0144005.	PR 08-OCT-1999;	99US-0158232.
PR 16-JUL-1999;	99US-0144085.	PR 12-OCT-1999;	99US-0158369.
PR 16-JUL-1999;	99US-0144086.	PR 13-OCT-1999;	99US-0159293.
PR 19-JUL-1999;	99US-0144325.	PR 13-OCT-1999;	99US-0159294.
PR 19-JUL-1999;	99US-0144331.	PR 14-OCT-1999;	99US-0159329.
PR 19-JUL-1999;	99US-0144332.	PR 14-OCT-1999;	99US-0159330.
PR 19-JUL-1999;	99US-0144333.	PR 14-OCT-1999;	99US-0159331.
PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144352.	PR 18-OCT-1999;	99US-0159638.
PR 20-JUL-1999;	99US-0144884.	PR 21-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-014814.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0148086.	PR 21-OCT-1999;	99US-0160767.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160815.
PR 22-JUL-1999;	99US-0145192.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145145.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145218.	PR 22-OCT-1999;	99US-0160989.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 26-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 04-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		

Query Match 76.7%; Score 33; DB 21; Length 175;
Best Local Similarity 66.7%; Pred. No. 86;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 GLIEKNIEL 9
Db 94 GLISQVEL 102
RESULT 10
AAG17121
ID AAG17121 standard; Protein; 203 AA.
XX

AC AAG17121;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18023.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0137282.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0136094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 23-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

```
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161408.
PR 25-OCT-1999; 99US-0161409.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.7%; Score 33; DB 21; Length 203;
Best Local Similarity 66.7%; Pred. No. le+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIERNIEL 9
Db 122 GLISQVEL 130
||| :|||

RESULT 11
AAAG17120
ID AAG17120 standard; Protein; 306 AA.
XX
AC AAG17120;
XX
DF 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 18022.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
```

```
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
```

PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.7%; Score 33; DB 21; Length 306;
Best Local Similarity 66.7%; Pred. NO. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
||| :|:|
Db 225 GLISQVEL 233

RESULT 12

AAW55960
ID AAW55960 standard; Protein; 974 AA.

XX AAW55960;

XX AC
XX 28-AUG-1998 (first entry)

XX Human transient receptor potential protein Htrp1.

XX Htrp1; transient receptor potential; trp protein; human;
KW capacitative calcium ion entry; CCE; asthma; hypertension;
KW diabetes; osteoporosis; osteogenesis; thrombosis; immunodeficiency;
KW gene therapy.

XX Homo sapiens.

XX OS

XX WO9808979-A1.

PN

XX PD 05-MAR-1998.
XX PF 29-AUG-1997; 97WO-US15247.
XX PR 15-OCT-1996; 96US-0729955.
XX PR 30-AUG-1996; 96US-0025111.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Birnbaumer L, Zhu X;
XX PI WPI; 1998-230269/20.
XX DR N-PSDB; AAV26029.
XX
XX Controlling capacitative calcium ion entry into mammalian cells - by
PT changing activity of transient receptor potential proteins, e.g. for
PT treating asthma, hypertension etc.
XX
XX Claim 9; Page 27-30; 60pp; English.
XX
XX This polypeptide comprises a human transient receptor potential
CC (trp) protein, designated Htrp1, that is an essential part of the
CC capacitative calcium ion entry (CCE) mechanism in human cells.
CC Htrp1 cDNA (see AAV26029) was isolated from a kidney cDNA library
CC following a database search for sequences homologous to the deduced
CC amino acid sequence of Drosophila trp. A fragment of Htrp1 cDNA
CC was used to show expression of the Htrp1 mRNA in a variety of human
CC tissues. Htrp3 (see AAW55961) has also been identified. CCE into
CC a mammalian cell expressing a trp protein required for CCE is
CC controlled in a claimed method by treating the cell with an agent
CC that increases or decreases the amount of biologically active trp
CC protein from its normal level. Agents that inhibit CCE are
CC potentially useful for treating asthma, hypertension and
CC osteoporosis, also for antithrombotic therapy, while those that
CC stimulate CCE are used to treat type II diabetes and to induce bone
CC formation. Primary immunodeficiency, if associated with trp gene
CC mutations, may be treated by gene therapy.
XX
XX Sequence 974 AA;
XX
XX Query Match 76.7%; Score 33; DB 19; Length 974;
XX Best Local Similarity 66.7%; Pred. No. 5.5e+02;
XX Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GLIEKNIEL 9
XX I :I:I:I:I
XX Db 874 GALERNIEL 882
XX
XX RESULT 13
XX ABB69318
XX ID ABB69318 standard; Protein; 1056 AA.
XX AC ABB69318;
XX
XX DT 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster polypeptide SEQ ID NO 34746.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX

XX (PEKE) PE CORP NY.
XX PA Venter JC, Adams M, Li PWD, Myers EW;
XX PI WPI; 2001-656860/75.
XX DR N-PSDB; ABL13421.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX Disclosure; SEQ ID NO 34746; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 1056 AA;
XX
XX Query Match 76.7%; Score 33; DB 22; Length 1056;
XX Best Local Similarity 77.8%; Pred. No. 6e+02;
XX Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 GLIEKNIEL 9
XX I:I:I:I:I
XX Db 691 GLIEKEIKL 699
XX
XX RESULT 14
XX ABP38169
XX ID ABP38169 standard; Protein; 122 AA.
XX
XX AC ABP38169;
XX
XX DT 24-JUL-2002 (first entry)
XX
XX Staphylococcus epidermidis ORF amino acid sequence SEQ ID NO:3014.
XX
XX DE Staphylococcus epidermidis; open reading frame; ORF; bacterial infection;
XX KW antibacterial; gene therapy.
XX
XX OS Staphylococcus epidermidis.
XX
XX PN US6380370-B1.
XX
XX PD 30-APR-2002.
XX
XX PF 13-AUG-1998; 98US-0134001.
XX
XX PR 14-AUG-1997; 97US-055779P.
XX
XX PR 08-NOV-1997; 97US-064964P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2002-381255/41.
XX
XX N-PSDB; ABN90714.
XX
XX Novel isolated nucleic acid encoding a Staphylococcus epidermidis
PT polypeptide, useful for diagnosing and treating bacterial infections -
XX
XX Disclosure; SEQ ID 3014; 267pp; English.
XX

CC ABN90538 to ABN93374 represent Staphylococcus epidermidis open reading
 CC frame (ORF) nucleic acid sequences which encode the amino acid sequences
 CC given in ABP35124 to ABP37960. The S. epidermidis sequences have
 CC antibacterial activity and can be used in gene therapy. The sequences
 CC can also be used in the diagnosis and treatment of bacterial infections,
 CC particularly S. epidermidis infections. The sequences can be used to
 CC screen for compounds able to interfere with the S. epidermidis life
 CC cycle or inhibit S. epidermidis infection.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from the
 CC USPTO web site.
 XX
 SQ Sequence 122 AA;
 Query Match 74.4%; Score 32; DB 23; Length 122;
 Best Local Similarity 85.7%; Pred. No. 92;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 3 IEKNIEL 9
 Db 62 IEKNVEL 68
 ||||:|
 ||||:|
 RESULT 15
 AAU20516
 ID AAU20516 standard; Protein; 235 AA.
 XX
 AC AAU20516;
 DT 06-DEC-2001 (first entry)
 XX
 DE Human secreted protein, Seq ID No 508.
 XX
 KW Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;
 KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;
 KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;
 KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;
 KW multiple sclerosis; cancer; hyperproliferative disorder; infection;
 KW Gaucher's disease; neurological disease; cerebrovascular disorder;
 KW thrombosis; wound healing.
 XX
 OS Homo sapiens.
 XX
 PN WO200155326-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01347.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR WPI; 2001-451931/48.
 DR N-PSDB; AAS33225.
 XX
 PT New nucleic acids and polypeptides, useful for diagnosing, preventing
 PT or treating medical conditions -
 XX
 PS Claim 11; SEQ ID No 508; 753pp; English.
 XX
 CC The invention relates to novel isolated nucleic acid molecules (I)
 CC encoding human secreted proteins (II). (I) and (II) are used to prevent,
 CC treat or ameliorate a medical condition in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. (I) and (II) may be used in
 CC the prevention, treatment and diagnosis of diseases associated with
 CC inappropriate expression of secreted proteins. (I) and complementary
 CC sequences may also be used as DNA probes in diagnostic assays (e.g.
 CC polymerase chain reactions (PCR)) to detect and quantitate the presence
 CC of similar nucleic acid sequences in samples, and so which patients may
 CC be in need of restorative therapy. (II) may also be used as antigens in

CC the production of antibodies and in assays to identify modulators
 CC (agonists and antagonists) of the expression and activity of the secreted
 CC proteins. The anti-(II) antibodies and antagonists may also be used to
 CC down regulate expression and activity of (II). The anti-(II) antibodies
 CC may also be used as diagnostic agents for detecting the presence of (II)
 CC in samples (e.g. by enzyme linked immunosorbant assay (ELISA)). The
 CC disorders include for example: immune/autoimmune diseases (e.g. HIV
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,
 CC parkinson's disease and Charcot-Marie-Tooth disease), cardio-/
 CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia,
 CC angina and thrombosis), infections caused by bacteria, viruses and
 CC fungi and ocular disorders (e.g. corneal infections). (I) and (II),
 CC agonists, antagonists and antibodies can also be used to promote wound
 CC healing, maintain organs before transplantation, and support cell culture
 CC of primary tissues. AAU20342-AAU20666 represent human secreted protein
 CC amino acid sequences, and related sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification but was obtained in electronic format directly from WIPO
 CC at: ftp.wipo.int/pub/published_pct_sequences.
 XX

SQ Sequence 235 AA;

Query Match 74.4%; Score 32; DB 22; Length 235;
 Best Local Similarity 55.6%; Pred. No. 1.9e+02;
 Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
 Db 222 GLIEKNIEL 230
 |:::|::|

Search completed: February 4, 2003, 17:38:37
 Job time : 37 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:36:38 ; Search time 15 seconds
(without alignments)
17.654 Million cell updates/sec

Title: US-09-865-548A-13
Perfect score: 43
Sequence: 1 GLIEKNIEL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32	74.4	122	4	US-09-134-001C-3014
2	31	72.1	428	4	US-09-134-001C-2942
3	30	69.8	53	2	US-08-676-782-9
4	30	69.8	266	4	US-09-134-001C-4962
5	30	69.8	394	4	US-09-252-816A-1
6	30	69.8	504	1	US-08-441-139-18
7	30	69.8	521	2	US-08-557-122A-32
8	30	69.8	521	4	US-09-262-666-32
9	30	69.8	522	4	US-09-368-588-2
10	30	69.8	530	2	US-08-557-122A-35
11	30	69.8	530	4	US-09-262-666-35
12	30	69.8	728	4	US-09-298-924-2
13	30	69.8	3052	2	US-08-557-122A-26
14	30	69.8	3052	4	US-09-262-666-26
15	29	67.4	194	4	US-09-218-363-17
16	29	67.4	255	1	US-08-622-353-8
17	29	67.4	255	2	US-08-622-352A-10
18	29	67.4	255	3	US-08-826-390-10
19	29	67.4	256	2	US-08-211-312-5
20	29	67.4	256	3	US-08-472-285-5
21	29	67.4	256	4	US-08-472-929-5
22	29	67.4	256	4	US-09-305-489-2
23	29	67.4	266	4	US-09-218-363-8
24	29	67.4	267	4	US-09-218-363-10
25	29	67.4	267	4	US-09-218-363-4
26	29	67.4	329	4	US-09-424-349A-8
27	29	67.4	380	1	US-08-585-758A-4

```

28 29 67.4 380 1 US-08-977-818-4 Sequence 4, Appli
29 29 67.4 380 2 US-08-670-274B-4 Sequence 4, Appli
30 29 67.4 380 4 US-09-146-187-4 Sequence 4, Appli
31 29 67.4 381 1 US-08-585-758A-2 Sequence 2, Appli
32 29 67.4 381 1 US-08-977-818-2 Sequence 2, Appli
33 29 67.4 381 2 US-08-670-274B-2 Sequence 2, Appli
34 29 67.4 381 2 US-08-786-999-3 Sequence 3, Appli
35 29 67.4 381 4 US-09-146-187-2 Sequence 1, Appli
36 29 67.4 390 2 US-08-786-999-1 Sequence 2, Appli
37 29 67.4 410 2 US-08-723-415B-10 Sequence 10, Appli
38 29 67.4 410 4 US-09-189-627A-10 Sequence 10, Appli
39 29 67.4 410 4 US-09-710-861-10 Sequence 10, Appli
40 29 67.4 1861 2 US-08-790-912-4 Sequence 4, Appli
41 28 65.1 30 1 US-08-248-505-5 Sequence 5, Appli
42 28 65.1 46 1 US-08-195-152-4 Sequence 4, Appli
43 28 65.1 50 2 US-08-980-071-44 Sequence 44, Appli
44 28 65.1 50 2 US-08-757-536-44 Sequence 44, Appli
45 28 65.1 50 3 US-09-314-093-44 Sequence 44, Appli

```

ALIGNMENTS

```

RESULT 1
US-09-134-001C-3014
; Sequence 3014, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3014
; LENGTH: 122
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3014

```

```

Query Match 74.4%; Score 32; DB 4; Length 122;
Best Local Similarity 85.7%; Pred. No. 24;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 IEKNIEL 9
Db 62 IEKNVEL 68
|||||

```

```

RESULT 2
US-09-134-001C-2942
; Sequence 2942, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2942
; LENGTH: 428
; TYPE: PRT

```

; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2942

Query Match 72.1%; Score 31; DB 4; Length 428;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIE 8
 |||||
Db 200 LIEKNLE 206

RESULT 3

US-08-676-782-9
; Sequence 9, Application US/08676782
; Patent No. 5976792
; GENERAL INFORMATION:
; APPLICANT: CHEUNG, Ambrose
; APPLICANT: FISCHETTI, Vincent A.
; TITLE OF INVENTION: REGULATION OF EXOPROTEIN IN
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,782
; FILING DATE: 08-JUL-1996
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/248,505
; FILING DATE: 25-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 016921-092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021

; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 53 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-676-782-9

Query Match 69.8%; Score 30; DB 2; Length 53;
Best Local Similarity 71.4%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIE 8
 |||||
Db 23 IVEKNIE 29

RESULT 4

US-09-134-001C-4962
; Sequence 4962, Application US/09134001C
; Patent No. 6360370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS

; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13

; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 4962
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-4962

Query Match 69.8%; Score 30; DB 4; Length 266;
Best Local Similarity 62.5%; Pred. No. 1.4e+02;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
 |||||
Db 117 LIEKNVKI 124

RESULT 5

US-09-252-816A-1
; Sequence 1, Application US/09252816A
; Patent No. 6265633
; GENERAL INFORMATION:
; APPLICANT: OKADA, Yukio

; APPLICANT: ITO, Kazutoshi
; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE
; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE
; TITLE OF INVENTION: SAME
; FILE REFERENCE: 1959-0008-0
; CURRENT APPLICATION NUMBER: US/09/252,816A
; CURRENT FILING DATE: 1999-02-19
; PRIOR APPLICATION NUMBER: JP HEI 10-37266
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: JP HEI 10-174235
; PRIOR FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Humulus lupulus
US-09-252-816A-1

Query Match 69.8%; Score 30; DB 4; Length 394;
Best Local Similarity 75.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
 |||
Db 276 GLISQNI 283

RESULT 6

US-08-441-139-18
; Sequence 18, Application US/08441139
; Patent No. 5773245
; GENERAL INFORMATION:
; APPLICANT: Wittrup, Dr. Karl D.

; APPLICANT: Robinson, Anne S.
; TITLE OF INVENTION: METHODS FOR INCREASING SECRETION OF
; TITLE OF INVENTION: RECOMBINANTLY EXPRESSED PROTEINS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: NY

; COUNTRY: USA
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/441,139
; FILING DATE: 15-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/089,997
; FILING DATE: 06-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DIGILLO, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 8646
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 516-742-4343
; TELEFAX: 516-742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-441-139-18

Query Match 69.8%; Score 30; DB 1; Length 504;
Best Local Similarity 75.0%; Pred No. 2.9e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
; : : : : :
; : : : : :
Db 58 LVEKNITL 65

RESULT 7
US-08-557-122A-32
; Sequence 32, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,122A
; FILING DATE: 11-DEC-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-557-122A-32

Query Match 69.8%; Score 30; DB 2; Length 521;
Best Local Similarity 75.0%; Pred No. 3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
; : : : : :
; : : : : :
Db 78 LVEKNITL 85

RESULT 8
US-09-262-666-32
; Sequence 32, Application US/09262666
; Patent No. 6346244
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6346244o No. 6346244disk of No. 6346244th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 521 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-262-666-32

Query Match 69.8%; Score 30; DB 4; Length 521;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
; : : : : :
; : : : : :
Db 78 LVEKNITL 85

RESULT 9
US-09-368-588-2
; Sequence 2, Application US/09368588
; Patent No. 6387683
; GENERAL INFORMATION:
; APPLICANT: ISHII, NOBUYOSHI

APPLICANT: SUZUKI, YASUO
APPLICANT: UCHIDA, KOHJI
APPLICANT: MATUO, YUSHI
APPLICANT: TANAKA, HIDEO
TITLE OF INVENTION: RECOMBINANT YEAST PDI AND PROCESS FOR PRODUCTION THEREOF
FILE REFERENCE: 139-32
CURRENT APPLICATION NUMBER: US/09/368,588
CURRENT FILING DATE: 1999-08-05
PRIOR APPLICATION NUMBER: PCT/JP98/00498
PRIOR FILING DATE: 1998-02-06
PRIOR APPLICATION NUMBER: JP 38588/1997
PRIOR FILING DATE: 1997-02-07
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 522
TYPE: PRT
ORGANISM: Saccharomyces cervisiae
US-09-368-588-2

Query Match 69.8%; Score 30; DB 4; Length 522;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
I:|||||
Db 78 LVEKNITL 85

RESULT 10
US-08-557-122A-35
Sequence 35, Application US/08557122A
Patent No. 5879664
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Mailand
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5879664o No. 5879664disk of No. 5879664th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9635
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-557-122A-35

Query Match 69.8%; Score 30; DB 2; Length 530;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
I:|||||
Db 78 LVEKNITL 85
RESULT 11
US-09-262-666-35
Sequence 35, Application US/09262666
Patent No. 6346244
GENERAL INFORMATION:
APPLICANT: Hjort, Carsten Mailand
TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6346244o No. 6346244disk of No. 6346244th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/262,666
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,122
FILING DATE: 11-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9635
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-262-666-35

Query Match 69.8%; Score 30; DB 4; Length 530;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
I:|||||
Db 78 LVEKNITL 85

RESULT 12
US-09-298-924-2
Sequence 2, Application US/09298924
Patent No. 6391595
GENERAL INFORMATION:
APPLICANT: KATO, Masaru
MIURA, Yutaka
KETTORU, Masako
IWAMATSU, Akihiro
KOBAYASHI, Kazuo
KOMEDA, Toshihiro
TITLE OF INVENTION: NOVEL TRANSFERASE AND AMYLASE, PROCESS
FOR PRODUCING THE ENZYMES, USE THEREOF, AND GENE CODING
FOR THE SAME
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:

ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/298,924
FILING DATE: 26-Apr-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/750,569
FILING DATE: <Unknown>
APPLICATION NUMBER: JP 7-120673
FILING DATE: 21-APR-1995
APPLICATION NUMBER: JP 6-311185
FILING DATE: 21-NOV-1994
APPLICATION NUMBER: JP 6-286917
FILING DATE: 21-NOV-1994
APPLICATION NUMBER: JP 6-290394
FILING DATE: 31-OCT-1994
APPLICATION NUMBER: JP 6-194223
FILING DATE: 18-AUG-1994
APPLICATION NUMBER: JP 6-133354
FILING DATE: 16-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 49441/110
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 728 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-298-924-2

Query Match 69.8%; Score 30; DB 4; Length 728;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
|:|||||
Db 629 LIEKNIEL 636

RESULT 13
US-08-557-122A-26
; Sequence 26, Application US/08557122A
; Patent No. 5879664
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5879664 No. 5879664disk of No. 5879664th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,122A
FILING DATE: 11-DEC-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3980.204-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 3052 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-557-122A-26

Query Match 69.8%; Score 30; DB 2; Length 3052;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
|:|||||
Db 521 LVEKNITL 528

RESULT 14
US-09-262-666-26
; Sequence 26, Application US/09262666
; Patent No. 6346244
; GENERAL INFORMATION:
; APPLICANT: Hjort, Carsten Mailand
; TITLE OF INVENTION: Fungal Protein Disulfide Isomerase
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6346244 No. 6346244disk of No. 6346244th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/262,666
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/557,122
; FILING DATE: 11-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3980.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3052 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide

US-09-262-666-26

Query Match 69.8%; Score 30; DB 4; Length 3052;
Best Local Similarity 75.0%; Pred. No. 2.1e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
|:|||||
Db 521 LVEKNITL 528

RESULT 15

US-09-218-363-17
; Sequence 17, Application US/09218363
; Patent No. 6387616
; GENERAL INFORMATION:
; APPLICANT: Ozelius, Laurie J.
; APPLICANT: Breakfield, Xandra O.
; TITLE OF INVENTION: TORSIN, TORSIN GENES, AND METHODS OF USE
; FILE REFERENCE: MGH-1184PA2
; CURRENT APPLICATION NUMBER: US/09/218,363
; CURRENT FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 09/099,454
; EARLIER FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: 60/050,244
; EARLIER FILING DATE: 1997-06-19
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(194)
; OTHER INFORMATION: xaa = Any Amino Acid
US-09-218-363-17

Query Match 67.4%; Score 29; DB 4; Length 194;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNI 7
|||:|:
Db 165 GLIDKNL 171

Search completed: February 4, 2003, 17:38:59
Job time : 17 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:37:58 : Search time 11 Seconds
(without alignments)
18.139 Million cell updates/sec

Title: US-09-865-548A-13
Perfect score: 43
Sequence: 1 GLIEKNIEL 9

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 129505 seqs, 22169297 residues
Total number of hits satisfying chosen parameters: 129505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	34	79.1	254	10	US-09-815-242-10730
2	34	79.1	360	9	US-09-829-378-3
3	32	74.4	1161	9	US-10-170-102-4
4	30	69.8	394	10	US-09-837-654-1
5	30	69.8	394	10	US-09-837-554-1
6	30	69.8	637	10	US-09-815-242-12058
7	30	69.8	950	10	US-09-815-242-11961
8	29	67.4	168	10	US-09-925-637-68
9	29	67.4	192	10	US-09-925-302-728
10	29	67.4	193	10	US-09-811-284-164
11	29	67.4	194	10	US-09-772-105-17
12	29	67.4	256	10	US-09-815-242-5563
13	29	67.4	262	10	US-09-815-242-5072
14	29	67.4	266	10	US-09-772-105-8
15	29	67.4	267	10	US-09-772-105-10
16	29	67.4	267	10	US-09-815-242-12318
17	29	67.4	267	10	US-09-815-242-13099
18	29	67.4	267	10	US-09-815-242-105-4
19	29	67.4	290	10	US-09-772-105-4

20	29	67.4	367	9	US-09-992-598-8	Sequence 8, Appli
21	29	67.4	367	9	US-09-989-293A-8	Sequence 8, Appli
22	29	67.4	367	9	US-09-989-735-8	Sequence 8, Appli
23	29	67.4	367	9	US-09-990-444-8	Sequence 8, Appli
24	29	67.4	367	9	US-09-989-730-8	Sequence 8, Appli
25	29	67.4	367	9	US-09-990-436-8	Sequence 8, Appli
26	29	67.4	367	9	US-09-991-181-8	Sequence 8, Appli
27	29	67.4	367	9	US-09-993-687-8	Sequence 8, Appli
28	29	67.4	367	9	US-09-989-734-8	Sequence 8, Appli
29	29	67.4	367	9	US-09-997-653-8	Sequence 8, Appli
30	29	67.4	367	9	US-09-993-667-8	Sequence 8, Appli
31	29	67.4	367	10	US-09-989-722-8	Sequence 8, Appli
32	29	67.4	367	10	US-09-989-723-8	Sequence 8, Appli
33	29	67.4	367	10	US-09-989-279-8	Sequence 8, Appli
34	29	67.4	367	10	US-09-989-727-8	Sequence 8, Appli
35	29	67.4	367	10	US-09-989-731-8	Sequence 8, Appli
36	29	67.4	367	10	US-09-989-732-8	Sequence 8, Appli
37	29	67.4	367	10	US-09-991-073-8	Sequence 8, Appli
38	29	67.4	367	10	US-09-990-442-8	Sequence 8, Appli
39	29	67.4	367	10	US-09-991-163-8	Sequence 8, Appli
40	29	67.4	367	10	US-09-993-604-8	Sequence 8, Appli
41	29	67.4	367	10	US-09-990-456-8	Sequence 8, Appli
42	29	67.4	367	10	US-09-989-721-8	Sequence 8, Appli
43	29	67.4	380	10	US-09-804-690-4	Sequence 4, Appli
44	29	67.4	381	10	US-09-804-690-2	Sequence 2, Appli
45	29	67.4	483	9	US-09-738-626-4842	Sequence 4842, Ap

ALIGNMENTS

RESULT 1
US-09-815-242-10730
: Sequence 10730, Application US/09815242
: Patent No. US20020061569A1
: GENERAL INFORMATION:
: APPLICANT: Haselbeck, Robert
: APPLICANT: Ohlsen, Karl L.
: APPLICANT: Zyskind, Judith W.
: APPLICANT: Wall, Daniel
: APPLICANT: Trawick, John D.
: APPLICANT: Carr, Grant J.
: APPLICANT: Yamamoto, Robert T.
: APPLICANT: Xu, H. Howard
: TITLE OF INVENTION: Identification of Essential Genes in
: FILE REFERENCE: ELITRA.011A
: CURRENT APPLICATION NUMBER: US/09/815,242
: CURRENT FILING DATE: 2001-03-21
: PRIOR APPLICATION NUMBER: 60/191,078
: PRIOR FILING DATE: 2000-03-21
: PRIOR APPLICATION NUMBER: 60/206,848
: PRIOR FILING DATE: 2000-03-23
: PRIOR APPLICATION NUMBER: 60/207,727
: PRIOR FILING DATE: 2000-03-26
: PRIOR APPLICATION NUMBER: 60/242,578
: PRIOR FILING DATE: 2000-10-23
: PRIOR APPLICATION NUMBER: 60/253,625
: PRIOR FILING DATE: 2000-11-27
: PRIOR APPLICATION NUMBER: 60/257,931
: PRIOR FILING DATE: 2000-12-22
: PRIOR APPLICATION NUMBER: 60/269,308
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 14110
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 10730
: LENGTH: 254
: TYPE: PRI
: ORGANISM: Enterococcus faecalis
US-09-815-242-10730

Query Match 79.1% Score 34: DB 10: Length 254;
Best Local Similarity 75.0%; Pred. No. 12;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
 |||||:|
 DB 167 LIEKNVEI 174

RESULT 2

US-09-829-378-3
 ; Sequence 3, Application US/09829378
 ; Patent No. US20020170082A1
 ; GENERAL INFORMATION:
 ; APPLICANT: FOX, Timothy
 ; APPLICANT: ALBERTSON, Marc C.
 ; TITLE OF INVENTION: GENE AFFECTING MALE FERTILITY IN PLANTS
 ; FILE REFERENCE: 033229/0631
 ; CURRENT APPLICATION NUMBER: US/09/829,378
 ; CURRENT FILING DATE: 2001-05-09
 ; PRIOR APPLICATION NUMBER: US 09/340,684
 ; PRIOR FILING DATE: 2001-06-01
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 360
 ; TYPE: PRT
 ; ORGANISM: Petunia hybrida
 ; FEATURE: VARIANT
 ; NAME/KEY: (5)
 ; LOCATION: (5)
 ; OTHER INFORMATION: xaa is any amino acid
 US-09-829-378-3

Query Match 79.1%; Score 34; DB 9; Length 360;
 Best Local Similarity 87.5%; Pred. No. 17;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
 |||||
 DB 289 GLISKNI 296

RESULT 3

US-10-170-102-4
 ; Sequence 4, Application US/10170102
 ; Publication No. US20030003539A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Curtis, Rory A.J.
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: 67109: A Human Phospholipid Transporter
 ; TITLE OF INVENTION: Family Member and Uses Therefor
 ; FILE REFERENCE: MF101-059P1RM
 ; CURRENT APPLICATION NUMBER: US/10/170,102
 ; CURRENT FILING DATE: 2002-06-12
 ; PRIOR APPLICATION NUMBER: 60/297840
 ; PRIOR FILING DATE: 2001-06-13
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 1161
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-170-102-4

Query Match 74.4%; Score 32; DB 9; Length 1161;
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
 |||||:|
 DB 643 LIEKNQL 650

RESULT 4

US-09-837-654-1
 ; Sequence 1, Application US/09837654
 ; Patent No. US2002010952A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ITO, Kazutoshi
 ; APPLICANT: OKADA, Yukio
 ; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE
 ; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE
 ; TITLE OF INVENTION: SAME
 ; FILE REFERENCE: 1959-0008-0
 ; CURRENT APPLICATION NUMBER: US/09/837,654
 ; CURRENT FILING DATE: 2001-04-19
 ; EARLIER APPLICATION NUMBER: 09/252,816
 ; EARLIER FILING DATE: 1999-02-19
 ; EARLIER APPLICATION NUMBER: JP HEI 10-174235
 ; EARLIER FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Humulus lupulus
 US-09-837-654-1

Query Match 69.8%; Score 30; DB 10; Length 394;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
 ||||:|
 DB 276 GLISQNI 283

RESULT 5

US-09-837-554-1
 ; Sequence 1, Application US/09837554
 ; Patent No. US20020102761A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ITO, Kazutoshi
 ; APPLICANT: OKADA, Yukio
 ; TITLE OF INVENTION: ISOLATED AND PURIFIED NUCLEIC ACIDS COMPRISING A GENE
 ; TITLE OF INVENTION: AND A REGULATORY REGION FOR THE GENE EXPRESSION OF THE
 ; TITLE OF INVENTION: SAME
 ; FILE REFERENCE: 1959-0008-0
 ; CURRENT APPLICATION NUMBER: US/09/837,554
 ; CURRENT FILING DATE: 2001-04-19
 ; PRIOR APPLICATION NUMBER: 09/252,816
 ; PRIOR FILING DATE: 1999-02-19
 ; PRIOR APPLICATION NUMBER: JP HEI 10-37266
 ; PRIOR FILING DATE: 1998-02-19
 ; PRIOR APPLICATION NUMBER: JP HEI 10-174235
 ; PRIOR FILING DATE: 1998-06-22
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 394
 ; TYPE: PRT
 ; ORGANISM: Humulus lupulus
 US-09-837-554-1

Query Match 69.8%; Score 30; DB 10; Length 394;
 Best Local Similarity 75.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
 ||||:|
 DB 276 GLISQNI 283

RESULT 6

US-09-815-242-12058
 ; Sequence 12058, Application US/09815242
 ; Patent No. US20020061569A1

```
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12058
; LENGTH: 637
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12058

Query Match      69.8%; Score 30; DB 10; Length 637;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLIEKN 6
    |||||
Db 410 GLIEKN 415

RESULT 7
US-09-815-242-11961
; Sequence 11961, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12058
; LENGTH: 637
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12058
```

```
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 1410
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12058
; LENGTH: 637
; TYPE: PRF
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-11961

Query Match      69.8%; Score 30; DB 10; Length 950;
Best Local Similarity 66.7%; Pred. No. 3.2e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIEL 9
    |||||: ||
Db 880 GLIDKSAEL 888

RESULT 8
US-09-925-637-68
; Sequence 68, Application US/09925637
; Patent No. US2002010338A1
; GENERAL INFORMATION:
; APPLICANT: Choi
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Polypeptides
; FILE REFERENCE: PB560
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/23773
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/151,933
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 08/781,986
; PRIOR FILING DATE: 1997-01-03
; PRIOR APPLICATION NUMBER: US 08/956,171
; PRIOR FILING DATE: 1997-10-20
; PRIOR APPLICATION NUMBER: US 60/009,861
; PRIOR FILING DATE: 1996-01-06
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 68
; LENGTH: 168
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (47)..(47)
; OTHER INFORMATION: Xaa equals any amino acid
US-09-925-637-68

Query Match      67.4%; Score 29; DB 10; Length 168;
Best Local Similarity 62.5%; Pred. No. 71;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LIEKNIEL 9
    |||||:::
Db 19 LIEKNVAV 26

RESULT 9
US-09-925-302-728
; Sequence 728, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US/09/925,302
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
```

; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 728
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-302-728

Query Match 57.4%; Score 29; DB 10; Length 192;
Best Local Similarity 55.8%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Caps 0;

Qy 1 GLIEKNIE 9
||| |::|||
Db 152 GLTERDVEL 160

RESULT 10
US-09-811-284-164
; Sequence 164, Application US/09811284
; Patent No. US20020058306A1
; GENERAL INFORMATION:
; APPLICANT: Vogeli, Gabriel
; TITLE OF INVENTION: No. US20020058306A1el G Protein-Coupled Receptors
; FILE REFERENCE: 00167US1
; CURRENT FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/189,783
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,907
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,918
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,960
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/189,917
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/192,945
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,916
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,923
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,933
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,830
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/192,234
; PRIOR FILING DATE: 2000-03-27
; PRIOR APPLICATION NUMBER: 60/192,155
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,935
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 258
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 164
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-811-284-164

Query Match 57.4%; Score 29; DB 10; Length 193;
Best Local Similarity 75.0%; Pred. No. 83;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Caps 0;

Qy 1 GLIEKNIE 8
||| |::|||
Db 1 GLIEWNLE 8

RESULT 11
US-09-772-105-17
; Sequence 17, Application US/09772105
; Patent No. US20010029015A1
; GENERAL INFORMATION:
; APPLICANT: Ozeilius, Laurie J.
; TITLE OF INVENTION: TORSIN, TORSIN-RELATED GENES, AND
; FILE REFERENCE: 0838-1001009
; CURRENT APPLICATION NUMBER: US/09/772,105
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 09/218,363
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/099,454
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: US 60/050,244
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 194
; TYPE: PRT
; ORGANISM: Homo sapien
; NAME/KEY: VARIANT
; FEATURE:
; LOCATION: (1)...(194)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-772-105-17

Query Match 67.4%; Score 29; DB 10; Length 194;
Best Local Similarity 71.4%; Pred. No. 84;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLIEKNI 7
||| |::|||
Db 165 GLIDKNI 171

RESULT 12
US-09-815-242-5563
; Sequence 5563, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0


```
; SEQ ID NO 5563
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5563

Query Match
Best Local Similarity 67.4%; Score 29; DB 10; Length 248;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
Db 102 LIEKNVKV 109

RESULT 13
US-09-949-584-2
; Sequence 2, Application US/09949584
; Patent No. US20020119512A1
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: Jiang, Xinhua
; APPLICANT: Van Horn, Stephanie
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: ups
; FILE REFERENCE: GM10216
; CURRENT APPLICATION NUMBER: US/09/949,584
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US/09/305,489
; PRIOR FILING DATE: 1999-05-05
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-949-584-2

Query Match
Best Local Similarity 67.4%; Score 29; DB 10; Length 256;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
Db 107 LIEKNVKV 114

RESULT 14
US-09-815-242-5072
; Sequence 5072, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
```

```
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5072
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5072

Query Match
Best Local Similarity 67.4%; Score 29; DB 10; Length 262;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
Db 44 LLEKNIEV 51

RESULT 15
US-09-772-105-8
; Sequence 8, Application US/09772105
; Patent No. US20010029015A1
; GENERAL INFORMATION:
; APPLICANT: Ozellus, Laurie J.
; APPLICANT: Breakfield, Xandra O.
; TITLE OF INVENTION: TORSIN, TORSIN-RELATED GENES, AND
; TITLE OF INVENTION: METHODS OF DETECTING NEURONAL DISEASES
; FILE REFERENCE: 0838.1001009
; CURRENT APPLICATION NUMBER: US/09/772,105
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 09/218,363
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/099,454
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: US 60/050,244
; PRIOR FILING DATE: 1997-06-19
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 266
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-772-105-8

Query Match
Best Local Similarity 67.4%; Score 29; DB 10; Length 266;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNI 7
Db 193 GLIDKNL 199

Search completed: February 4, 2003, 17:42:07
Job time : 12 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:36:35 ; Search time 15 Seconds
(without alignments)
57.681 Million cell updates/sec

Title: US-09-865-548A-13

Perfect score: 43

Sequence: 1 GLIEKNIEL 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_73: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	100.0	1495	2 S22610	DNA (cytosine-5-)-
2	42	97.7	1490	2 JC5145	DNA (cytosine-5-)-
3	42	97.7	1573	2 S01845	DNA (cytosine-5-)-
4	42	97.7	1622	2 JE0378	DNA (cytosine-5-)-
5	40	93.0	1537	2 JC4172	DNA (cytosine-5-)-
6	35	81.4	265	2 C90387	hypothetical prote
7	35	81.4	408	2 B90517	conserved hypothet
8	35	81.4	592	2 S54489	phosphoribosylamin
9	34	79.1	232	2 S29556	naringenin-chalcon
10	34	79.1	323	2 E71134	naringenin-chalcon
11	34	79.1	381	2 S12223	hypothetical prote
12	34	79.1	388	1 SYSYCN	naringenin-chalcon
13	34	79.1	388	1 SYSYCL	naringenin-chalcon
14	34	79.1	388	1 SYSYC3	naringenin-chalcon
15	34	79.1	388	2 JQ2249	naringenin-chalcon
16	34	79.1	388	2 S60472	naringenin-chalcon
17	34	79.1	389	1 SYPJCN	naringenin-chalcon
18	34	79.1	389	1 SVPJCA	naringenin-chalcon
19	34	79.1	389	2 S49203	naringenin-chalcon
20	34	79.1	389	2 JC5136	naringenin-chalcon
21	34	79.1	390	1 SYSKCD	naringenin-chalcon
22	34	79.1	394	1 SYJCCS	naringenin-chalcon
23	34	79.1	395	1 SYISCL	naringenin-chalcon
24	34	79.1	395	1 SYISC3	naringenin-chalcon
25	34	79.1	396	2 S20515	naringenin-chalcon
26	34	79.1	398	2 S16275	naringenin-chalcon
27	34	79.1	398	2 S42523	naringenin-chalcon
28	34	79.1	398	2 S38190	naringenin-chalcon
29	34	79.1	398	2 S56699	naringenin-chalcon

30 34 79.1 400 1 SYZMW1 naringenin-chalcon
31 34 79.1 410 2 SL2224 naringenin-chalcon
32 34 79.1 1612 2 JC5210 DNA (cytosine-5-)-
33 33 76.7 325 1 F71066 hypothetical prote
34 33 76.7 405 2 A72383 sensor histidine k
35 33 76.7 462 2 B84680 probable salt-indu
36 33 76.7 605 2 D82434 probable conserved
37 33 76.7 785 2 T01025 hypothetical prote
38 33 76.7 1263 2 T13465 hypothetical prote
39 33 76.7 1790 2 S67593 transport protein
40 33 76.7 2225 2 T26063 hypothetical prote
41 32 74.4 521 2 C82922 methionyl-tRNA syn
42 32 74.4 594 2 T05544 hypothetical prote
43 32 74.4 610 2 H71612 asparagine-tRNA li
44 32 74.4 664 2 A97222 membrane associate
45 32 74.4 677 2 B82870 DNA topoisomerase

ALIGNMENTS

RESULT 1

S22610

DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) - human

C:Species: Homo sapiens (man)

C:Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 05-May-2000

C:Accession: S22610; S19183

R:Yen, R.W.C.; Vertino, P.M.; Nelkin, B.D.; Yu, J.J.; El-Deiry, W.; Kumaraswamy, A.;

Nucleic Acids Res. 20, 2287-2291, 1992

A:Title: Isolation and characterization of the cDNA encoding human DNA methyltransferase.

A:Reference number: S22610; MUID:92279022; PMID:1594447

A:Accession: S22610

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1495 <YEN>

A:Cross-references: EMBL:X63692

C:Keywords: DNA binding; methyltransferase; S-adenosylmethionine

Query Match 100.0%; Score 43; DB 2; Length 1495;

Best Local Similarity 100.0%; Pred. No. 3.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLIEKNIEL 9

Db 304 GLIEKNIEL 312

RESULT 2

JC5145

DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C:Date: 31-Jan-1997 #sequence_revision 27-Feb-1997 #text_change 21-Jul-2000

C:Accession: JC5145

R:Kimura, H.; Ishihara, G.; Tajima, S.

J. Biochem. 120, 1182-1189, 1996

A:Title: Isolation and expression of a Xenopus laevis DNA methyltransferase cDNA.

A:Reference number: JC5145; MUID:97164021; PMID:9010768

A:Contents: cocyte

A:Accession: JC5145

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1490 <KIM>

A:Cross-references: DDBJ:D78638; NID:g1731731; PIDN:BAAL1458.1; PID:g1731732

C:Comment: This enzyme is responsible for maintaining the methylation pattern once fo

C:Keywords: methyltransferase; S-adenosylmethionine

F:526-564/Region: cysteine-rich

F:983-995/Region: two-residue repeat (K-G)

Query Match

Best Local Similarity 97.7%; Score 42; DB 2; Length 1490;

Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GLIEKNIEL 9

Db 300 GLIEKNVEL 308
|||||:|

RESULT 3

S01845
DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) - mouse
N:Alternate names: DNA methyltransferase
C:Species: Mus musculus (house mouse)
C:Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 05-May-2000
C:Accession: S01845; A30461; S16605; S22719
R:Bestor, T.; Laudano, A.; Mattaliano, R.; Ingram, V.
J. Mol. Biol. 203, 971-983, 1998
A:Title: Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells.
A:Reference number: S01845; MUID:89094873; PMID:3210246
A:Accession: S01845
A:Molecule type: mRNA
A:Residues: 1-1573 <BES1>
A:Cross-references: EMBL:X14805
A:Accession: A30461
A:Molecule type: protein
A:Residues: 490-527 <BES2>
R:Bestor, T.
Submitted to the EMBL Data Library, March 1989
A:Reference number: S16605
A:Accession: S16605
A:Molecule type: mRNA
A:Residues: 1-1194, 'G', 1196-1573 <BES3>
A:Cross-references: EMBL:X14805
R:Bestor, T.H.

EMBO J. 11, 2611-2617, 1992
A:Title: Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regul
A:Reference number: S22719; MUID:92331613; PMID:1628623
A:Accession: S22719
A:Molecule type: protein
A:Residues: 986-1032 <BES4>
C:Keywords: DNA binding; methyltransferase; S-adenosylmethionine

Query Match 97.7%; Score 42; DB 2; Length 1573;
Best Local Similarity 88.9%; Pred. No. 5.2;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9

Db 312 GLIEKNVEL 320
|||||:|

RESULT 4

JE0378
DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 21-Jul-2000
C:Accession: JE0378
R:Kimura, H.; Takeda, T.; Tanaka, S.; Ogawa, T.; Shiota, K.
Biochem. Biophys. Res. Commun. 253, 495-501, 1998
A:Title: Expression of rat DNA (cytosine-5) methyltransferase (DNA MTase) in rodent troph
A:Reference number: JE0378; MUID:99097263; PMID:9878564
A:Accession: JE0378
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1622 <KIM>
A:Cross-references: DBJ:AB012214; NID:g4160669; PIDN:BAA37118.1; PID:g4160670
C:Keywords: methyltransferase; S-adenosylmethionine

Query Match 97.7%; Score 42; DB 2; Length 1622;
Best Local Similarity 88.9%; Pred. No. 5.4;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9

Db 431 GLIEKNVEL 439
|||||:|

RESULT 5

JC4172
DNA (cytosine-5-)-methyltransferase (EC 2.1.1.37) - chicken
C:Species: Gallus gallus (Chicken)
C:Date: 21-Nov-1995 #sequence_revision 08-Feb-1996 #text_change 21-Jul-2000
C:Accession: JC4172; PC4044
R:Tajima, S.; Tsuda, H.; Wakabayashi, N.; Asano, A.; Mizuno, S.; Nishimori, K.
J. Biochem. 117, 1050-1057, 1995
A:Title: Isolation and expression of a chicken DNA methyltransferase cDNA.
A:Reference number: JC4172; MUID:96172572; PMID:8586618
A:Accession: JC4172
A:Molecule type: mRNA
A:Residues: 1-1537 <TAJ>
A:Cross-references: DBJ:D43920; NID:g1374774; PIDN:BAA07867.1; PID:g1109610
A:Accession: PC4044
A:Molecule type: protein
A:Residues: 1055-1075; 1078-1097; 1131-1149; 1174-1195; 1205-1210; 1213-1232; 1459-1477; 1488-1490
C:Comment: This enzyme is a maintenance-type methylase that functions during DNA repl
C:Genetics:
A:Gene: cmt
C:Keywords: methyltransferase; S-adenosylmethionine
F:565-603/Region: cysteine-rich
F:1020-1033/Region: glycine/lysine-rich repeats

Query Match 93.0%; Score 40; DB 2; Length 1537;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9

Db 336 GLIEKNIEL 344
|||||:|

RESULT 6

C90387
hypothetical protein SSO2179 [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: C90387
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: C90387
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <KUR>
A:Cross-references: GB:AE006641; NID:g13815478; PIDN:AAK42354.1; GSPDB:GN00155
C:Genetics:
A:Gene: SSO2179

Query Match 81.4%; Score 35; DB 2; Length 265;
Best Local Similarity 75.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

Db 93 GILEKNIE 100
|:::|

RESULT 7

B90517
conserved hypothetical protein WYPU_0420 [imported] - Mycoplasma pulmonis (strain UAB
C:Species: Mycoplasma pulmonis
C:Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 03-Aug-2001
C:Accession: B90517
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer,
Nucleic Acids Res. 29, 2145-2153, 2001
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma p
A:Reference number: A99512; MUID:21267165; PMID:11353084
A:Accession: B90517

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-408 <KUR>
A:Cross-references: GB:AL445566; PID:g14089455; PIDN:CAC13215.1; GSPDB:GN00153
A:Experimental source: strain UAB CTIP
C:Genetics:
A:Gene: MYP0_0420
A:Genetic code: SGC3

Query Match 81.4%; Score 35; DB 2; Length 408;
Best Local Similarity 87.5%; Pred. No. 33;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 LIEKNIEL 9
:|||||
Db 230 LIEKNIEL 237

RESULT 8
S54489
phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.1.2.3) - yeast (Saccharomyces cerevisiae)
N:Alternate names: 5-aminimidazole-4-carboxamide ribotide transferase; protein YM858
C:Species: Saccharomyces cerevisiae
C:Date: 08-Jul-1995 #sequence_revision 19-Oct-1995 #text_change 05-Nov-1999
C:Accession: S54489
R:Lye, G.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54014
A:Accession: S54489
A:Molecule type: DNA
A:Residues: 1-592 <LYE>
A:Cross-references: EMBL:Z49273; NID:g809577; PIDN:CAA89269.1; PID:g809579; GSPDB:GN000153
C:Genetics:
A:Gene: SGD:ADE17; MIPS:YMR120c
A:Cross-references: SGD:S0004727; MIPS:YMR120c
A:Map position: 13R
C:Superfamily: purH bifunctional enzyme
C:Keywords: hydrolase; purine nucleotide biosynthesis; transferase

Query Match 81.4%; Score 35; DB 2; Length 592;
Best Local Similarity 66.7%; Pred. No. 49;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
|||||
Db 24 GLIEKNVRI 32

RESULT 9
S29556
naringenin-chalcone synthase (EC 2.3.1.74) - apple tree (fragment)
C:Species: Malus sp. (apple tree)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-May-2000
C:Accession: S29556
R:Podivinsky, E.; Bradley, J.M.; Davies, K.M.
submitted to the EMBL Data Library, October 1992
A:Reference number: S29554
A:Accession: S29556
A:Molecule type: mRNA
A:Residues: 1-232 <POD>
A:Cross-references: EMBL:X68977; NID:g19588; PIDN:CAA48773.1; PID:g19589
C:Superfamily: chalcone synthase
C:Keywords: acyltransferase; coenzyme A

Query Match 79.1%; Score 34; DB 2; Length 232;
Best Local Similarity 87.5%; Pred. No. 28;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
||| ||||
Db 116 GLISKNIIE 123

RESULT 10
E71134
hypothetical protein PH0841 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
C:Accession: E71134
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; S.M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogawa, Y.
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic bacterium
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71134
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-323 <KAW>
A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29935.1; PID:g3257252
A:Experimental source: strain ON3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0841
C:Superfamily: Methanobacterium thermoautotrophicum conserved hypothetical protein M

Query Match 79.1%; Score 34; DB 2; Length 323;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
|||||
Db 290 GLIEKSIE 297

RESULT 11
SI2223
naringenin-chalcone synthase (EC 2.3.1.74) 1 - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000
C:Accession: SI2223
R:O'Neill, S.D.; Tong, Y.; Spoerlein, B.; Forkmann, G.; Yoder, J.I.
Mol. Gen. Genet. 224, 279-288, 1990
A:Title: Molecular genetic analysis of chalcone synthase in Lycopersicon esculentum
A:Reference number: SI2223; MUID:91117196; PMID:1980524
A:Accession: SI2223
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-381 <ONE>
C:Genetics:
A:Map position: 9
C:Superfamily: chalcone synthase
C:Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 79.1%; Score 34; DB 2; Length 381;
Best Local Similarity 87.5%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
||| ||||
Db 268 GLISKNIIE 275

RESULT 12
SYSYN
naringenin-chalcone synthase (EC 2.3.1.74) 2 - soybean
C:Species: Glycine max (soybean)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 05-May-2000
C:Accession: SI6338; SI0475
R:Akada, S.; Kung, S.D.; Dube, S.K.
Nucleic Acids Res. 18, 3398, 1990
A:Title: Nucleotide sequence of one member of soybean chalcone synthase multi-gene family
A:Reference number: SI6338; MUID:90287722; PMID:2356130
A:Accession: SI6338
A:Molecule type: DNA
A:Residues: 1-388 <AKA>
A:Cross-references: EMBL:X52097; NID:g18751; PIDN:CAA36317.1; PID:g295803

C;Genetics:

A;Gene: chs

A;Introns: 60/1

C;Superfamily: chalcone synthase

C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match

Best Local Similarity 79.1%; Score 34; DB 1; Length 388;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

||| ||||

Db 272 GLISKNIIE 279

RESULT 13

SYSVC1

naringenin-chalcone synthase (EC 2.3.1.74) 1 - soybean

C;Species: Glycine max (soybean)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000

C;Accession: S15006

R;Akada, S.; Kung, S.D.; Dube, S.K.

Plant Mol. Biol. 16, 751-752, 1991

A;Title: The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene fa

A;Reference number: S15006; MUID:91329712; PMID:1868209

A;Accession: S15006

A;Molecule type: DNA

A;Residues: 1-388 <AKA>

A;Cross-references: EMBL:X54644; NID:g18561; PIDN:CRA38456.1; PID:g18562

C;Genetics:

A;Introns: 60/1

C;Superfamily: chalcone synthase

C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match

Best Local Similarity 79.1%; Score 34; DB 1; Length 388;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

||| ||||

Db 272 GLISKNIIE 279

RESULT 14

SYSYC3

naringenin-chalcone synthase (EC 2.3.1.74) 3 - soybean

C;Species: Glycine max (soybean)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 05-May-2000

C;Accession: S11486

R;Akada, S.; Kung, S.D.; Dube, S.K.

Nucleic Acids Res. 18, 5899, 1990

A;Title: The nucleotide sequence of gene 3 of the soybean chalcone synthase multigene fa

A;Reference number: S11486; MUID:91016949; PMID:2216793

A;Accession: S11486

A;Molecule type: DNA

A;Residues: 1-388 <AKA>

A;Cross-references: EMBL:X53958; NID:g18588; PIDN:CRA37909.1; PID:g18589

C;Genetics:

A;Introns: 60/1

C;Superfamily: chalcone synthase

C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match

Best Local Similarity 79.1%; Score 34; DB 1; Length 388;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

||| ||||

Db 272 GLISKNIIE 279

RESULT 15

JQ2249

naringenin-chalcone synthase (EC 2.3.1.74) - soybean

N;Alternate names: chalcone synthase

C;Species: Glycine max (soybean)

C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 05-May-2000

C;Accession: JQ2249; S21444

R;Akada, S.; Kung, S.D.; Dube, S.K.

Plant Physiol. 102, 317-319, 1993

A;Title: Nucleotide sequence and putative regulatory elements of gene 2 of the soybe

A;Reference number: JQ2249; MUID:94151428; PMID:8108500

A;Accession: JQ2249

A;Molecule type: DNA

A;Residues: 1-388 <AKA>

A;Cross-references: EMBL:X65636; NID:g18529; PIDN:CRA46590.1; PID:g18530

C;Comment: This enzyme catalyzes the condensation of one molecule of rho-coumaroyl-C

C;Superfamily: chalcone synthase

C;Keywords: acyltransferase; coenzyme A; flavonoid biosynthesis

Query Match 79.1%; Score 34; DB 2; Length 388;

Best Local Similarity 87.5%; Pred. No. 49;

Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

||| ||||

Db 272 GLISKNIIE 279

Search completed: February 4, 2003, 17:37:54

Job time : 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:36:34 ; Search time 10 Seconds
(without alignments)
37.329 Million cell updates/sec

Title: US-09-865-548A-13
Perfect score: 43
Sequence: 1 GLIEKNIEL 9

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43	100.0	1616	1	DNM1_HUMAN
2	42	97.7	1620	1	DNM1_MOUSE
3	42	97.7	1622	1	DNM1_RAT
4	40	93.0	1537	1	DNM1_CHICK
5	35	81.4	206	1	RR4_ASTLO
6	35	81.4	592	1	PU92_YEAST
7	34	79.1	232	1	CHSV_MALDO
8	34	79.1	388	1	CHSV_SOYBN
9	34	79.1	388	1	CHSV_SOYBN
10	34	79.1	388	1	CHSV_SOYBN
11	34	79.1	388	1	CHSV_SOYBN
12	34	79.1	389	1	CHSV_CMSI
13	34	79.1	389	1	CHSV_CASGL
14	34	79.1	389	1	CHSV_LYCES
15	34	79.1	389	1	CHSV_CMSI
16	34	79.1	389	1	CHSV_LYCES
17	34	79.1	389	1	CHSV_SOLTU
18	34	79.1	389	1	CHSV_PETHY
19	34	79.1	389	1	CHSV_SOLTU
20	34	79.1	389	1	CHSV_PEA
21	34	79.1	389	1	CHSV_SOLTU
22	34	79.1	389	1	CHSV_HYDMC
23	34	79.1	390	1	CHSV_ANTMA
24	34	79.1	391	1	CHSV_CITSI
25	34	79.1	391	1	CHSV_PEFER
26	34	79.1	392	1	CHSV_SECC
27	34	79.1	392	1	CHSV_PERAE
28	34	79.1	393	1	CHSV_VITVI
29	34	79.1	394	1	CHSV_SECC
30	34	79.1	394	1	CHSV_NATIN
31	34	79.1	394	1	CHSV_RAPSA
32	34	79.1	395	1	CHSV_SINAL
33	34	79.1	395	1	CHSV_SINAL

34	34	79.1	395	1	CHSV_PICMA
35	34	79.1	395	1	CHSV_PINST
36	34	79.1	396	1	CHSV_CHRAE
37	34	79.1	396	1	CHSV_PINSY
38	34	79.1	397	1	CHSV2_DAUCA
39	34	79.1	397	1	CHSV2_GERHY
40	34	79.1	398	1	CHSV1_HORVU
41	34	79.1	398	1	CHSV1_CALCH
42	34	79.1	398	1	CHSV1_ORYSA
43	34	79.1	398	1	CHSV1_PETCR
44	34	79.1	400	1	CHSV1_MAIZE
45	34	79.1	400	1	CHSV1_SORBI

Q9m5m0	picea maria
O65872	pinus strob
O04220	chrysosplen
P30079	pinus sylve
Q92840	daucus caro
P48390	gerbera hyb
P26018	hordeum vul
P48385	callistephu
P48405	oryza sativ
P16107	petroselinu
P24824	zea mays (m
Q9xgx1	sorghum dic

ALIGNMENTS

RESULT 1	
DNM1_HUMAN	
ID	DNM1_HUMAN STANDARD; PRT; 1616 AA.
AC	P26358; Q9UHG5; Q9UUA2; Q9UMZ6;
DT	01-MAY-1992 (Rel. 22, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	DNA (cytosine-5)-methyltransferase 1 (EC 2.1.1.37) (Dnmt1) (DNA methyltransferase Hsa1) (DNA Mase Hsa1) (MCMT) (M.Hsa1).
GN	DNMT1 OR DNMT OR AIM.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RP	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RX	MEDLINE-92279022; PubMed-1594447;
RA	Yen R.-W.C., Vertino P.M., Nelkin B.D., Yu J.J., Delry W.E., Kumaraswamy A., Lennon G.G., Trask B.J., Celano P., Baylin S.B.;
RA	"Isolation and characterization of the cDNA encoding human DNA methyltransferase."
RT	methyltransferase."
RL	Nucleic Acids Res. 20:2287-2291(1992).
RN	[2]
RP	REVISIONS TO N-TERMINUS.
RP	MEDLINE-97094871; PubMed-8940105;
RA	Yoder J.A., Yen R.-W.C., Vertino P.M., Bestor T.H., Baylin S.B.;
RT	"New 5' regions of the murine and human genes for DNA (cytosine-5)-methyltransferase."
RL	J. Biol. Chem. 271:31092-31097(1996).
RN	[3]
RP	SEQUENCE FROM N.A. (ISOFORM 3).
RC	TISSUE=Prostatic carcinoma;
RA	Li L.C., Au H., Chui R., Dahiya R.;
RT	"Human DNA methyltransferase (DNMT1) is alternatively spliced.;"
RL	Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN	[4]
RP	SEQUENCE OF 27-1087 FROM N.A. (ISOFORM 1).
RA	Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhart-Schultz K., Gordon L., Dias J., Ramirez M., Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A., Garnes J., Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andrest T., Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;
RT	"Sequence analysis of a 6 Mb region in 19p13.2 between D19S391 and D19S179.;"
RL	Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN	[5]
RP	PARTIAL SEQUENCE FROM N.A. (ISOFORM 2).
RX	MEDLINE-99380591; PubMed-10449766;
RA	Hsu D.-W., Lin M.-J., Lee T.-L., Wen S.-C., Chen X., Shen C.-K.J.;
RT	"Two major forms of DNA (cytosine-5) methyltransferase in human somatic tissues.;"
RL	Proc. Natl. Acad. Sci. U.S.A. 96:9751-9756(1999).
RN	[6]

J. Mol. Biol. 203:971-983(1988).
 [2]
 RN RP REVISIONS TO N-TERMINUS.
 RC
 RN TISSUE-Embryo;
 RP MEDLINE=97094871; PubMed=8940105;
 RX Yoder J.A., Yen R.-W.C., Vertino P.M., Bestor T.H., Baylin S.B.;
 RA "New 5' regions of the murine and human genes for DNA (cytosine-5)-
 RT methyltransferase.";
 RL J. Biol. Chem. 271:31092-31097(1996).
 [3]
 RN SEQUENCE FROM N.A. (ISOFORM 2).
 RP STRAIN=C57BL/6; TISSUE=Skeletal muscle;
 RX MEDLINE=20515133; PubMed=11063128;
 RA Aguirre-Arteta A.M., Grunewald I., Cardoso M.C., Leonhardt H.;
 RT "Expression of an alternative Dnmt1 isoform during muscle
 RL differentiation";
 RN Cell Growth Differ. 11:551-559(2000).
 [4]
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RP STRAIN=C57BL/6;
 RX STRAIN=20181859; PubMed=10715201;
 RA Margot J.B., Aguirre-Arteta A.M., Di Giacco B.V., Pradhan S.,
 RA Roberts R.J., Cardoso M.C., Leonhardt H.;
 RT "Structure and function of the mouse DNA methyltransferase gene: Dnmt1
 RL shows a tripartite structure.";
 RN J. Mol. Biol. 297:293-300(2000).
 [5]
 RN SEQUENCE OF 1-27 AND 119-1619 FROM N.A. (ISOFORMS 1 AND 2).
 RP MEDLINE=98119799; PubMed=9449671;
 RX Mertineit C., Yoder J.A., Taketo T., Laird D.W., Trasler J.M.,
 RA Bestor T.H.;
 RT "Sex-specific exons control DNA methyltransferase in mammalian germ
 RL cells.";
 RN Development 125:889-897(1998).
 [6]
 RN SEQUENCE OF 1-144 FROM N.A. (ISOFORMS 1 AND 2), AND SEQUENCE OF 3-6.
 RP STRAIN=129/Sv; and BALB/c; TISSUE=Embryonic stem cells;
 RX MEDLINE=99047652; PubMed=9830015;
 RA Gaudet F., Talbot D., Leonhardt H., Jaenisch R.;
 RT "A short DNA methyltransferase isoform restores methylation in vivo.";
 RL J. Biol. Chem. 273:32725-32729(1998).
 [7]
 RN SEQUENCE OF 1-119 FROM N.A. (ISOFORM 1).
 RP STRAIN=129/Sv; TISSUE=Embryonic stem cells, and Kidney;
 RX Tucker K.L., Talbot D., Lee M.A., Leonhardt H., Jaenisch R.;
 RA "Complementation of methylation deficiency in embryonic stem cells by
 RT a DNA methyltransferase minigene.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12920-12925(1996).
 [8]
 RN SEQUENCE OF 1-272 FROM N.A. (ISOFORM 1).
 RP STRAIN=C57BL/6J; TISSUE=Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staebli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RN Nature 403:685-690(2001).
 [9]
 RN PHOSPHORYLATION OF SER-515, AND MASS SPECTROMETRY.
 RP TISSUE=Erythroleukemia;
 RX MEDLINE=97362284; PubMed=9211941;
 RA Glickman J.F., Pavlovich J.G., Reich N.O.;
 RT "Peptide mapping of the murine DNA methyltransferase reveals a major
 RL phosphorylation site and the start of translation.";
 RN J. Biol. Chem. 272:117851-117857(1997).
 [10]
 RN INTERACTION WITH HDAC1.
 RP MEDLINE=20082816; PubMed=10615135;
 RX Fuks F., Burgers W.A., Brehm A., Hughes-Davies L., Kouzarides T.;
 RA "DNA methyltransferase Dnmt1 associates with histone deacetylase
 RT activity.";
 RL Nat. Genet. 24:88-91(2000).
 [11]
 RN INTERACTIONS WITH HDAC2 AND DNAP1.
 RP MEDLINE=20347709; PubMed=10888872;
 RX Rountree M.R., Bachman K.E., Baylin S.B.;
 RA "DNMT1 binds HDAC2 and a new co-repressor, DNAP1, to form a complex at
 RT replication foci.";
 RL Nat. Genet. 25:269-277(2000).
 [12]
 RN FUNCTION, AND SUBCELLULAR LOCATION.
 RP MEDLINE=21185930; PubMed=11290321;
 RX Howell C.Y., Bestor T.H., Ding F., Latham K.E., Mertineit C.,
 RA Trasler J.M., Chaillet J.R.;
 RT "Genomic imprinting disrupted by a maternal effect mutation in the
 RL Dnmt1 gene.";
 RN Cell 104:829-838(2001).
 [13]
 RN ALLOSTERIC ACTIVATION.
 RP MEDLINE=21293215; PubMed=11399088;
 RX Fatemi M., Hermann A., Pradhan S., Jeltsch A.;
 RA "The activity of the murine DNA methyltransferase Dnmt1 is controlled
 RT by interaction of the catalytic domain with the N-terminal part of
 RL the enzyme leading to an allosteric activation of the enzyme after
 RL binding to methylated DNA.";
 RN J. Mol. Biol. 309:1189-1199(2001).
 CC -1- FUNCTION: Methylates CpG residues. Preferentially methylates
 CC hemimethylated DNA. It is responsible for maintaining methylation
 CC patterns established in development. Isoform 2, in oocytes, may
 CC provide maintenance methyltransferase activity specifically at
 CC imprinted loci during the fourth embryonic S phase. Mediates
 CC transcriptional repression by direct binding to HDAC2.
 CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA = S-adenosyl-L-
 CC homocysteine + DNA containing 5-methylcytosine.
 CC -1- ENZYME REGULATION: Allosterically regulated. The binding of 5-
 CC methylcytosine-containing DNA to the N-terminal parts of Dnmt1
 CC causes an allosteric activation of the catalytic domain by a
 CC direct interaction of its Zn-binding domain with the catalytic
 CC domain.
 CC -1- SUBUNIT: Interacts with HDAC1 and with PCNA. Forms a complex with
 CC DNAP1 and HDAC2, with direct interaction.
 CC -1- SUBCELLULAR LOCATION: Nuclear; it is nucleoplasmic through most of
 CC the cell cycle and associates with replication foci during S-
 CC phase. In germ cells: spermatogonia, preleptotene and leptotene
 CC spermatocytes all express high levels of nuclear protein, while
 CC the protein is not detected in pachytene spermatocytes, despite
 CC the fact they expressed high levels of mRNA. In females, the
 CC protein is not detected in non-growing oocytes, in contrast to the
 CC growing oocytes. During the growing, the protein is no longer
 CC detectable in nuclei but accumulates to very high levels first
 CC throughout the cytoplasm. At the time of ovulation, all the
 CC protein is cytoplasmic and is actively associated with the oocyte
 CC cortex. After fecundation, in the preimplantation embryo, the
 CC protein remains cytoplasmic and after implantation, it is
 CC exclusively nuclear in all tissue types. Isoform 2 is sequestered
 CC in the cytoplasm of maturing oocytes and of preimplantation
 CC embryos, except for the 8-cell stage, while isoform 1 is
 CC exclusively nuclear.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: 1/long form (shown here) and
 CC 2/short form; are produced by alternative splicing.

```
CC -!- TISSUE SPECIFICITY: Isoform 1 is expressed in embryonic stem cells
CC and in somatic tissues. Isoform 2 is expressed in oocytes,
CC preimplantation embryos, testis and in skeletal muscle during
CC myogenesis.
CC
CC -!- DEVELOPMENTAL STAGE: In germ cells, it is present at high levels
CC in spermatogonia and spermatocytes until the pachytene stage,
CC where it falls to undetectable levels. The transient drop at the
CC pachytene stage coincides with the disappearance of the 5.2 kb
CC mRNA and the accumulation of a larger 6.0 kb mRNA. Oocytes
CC accumulate very large amounts of Dnmt1 protein during the growth
CC phase.
CC
CC -!- MISCELLANEOUS: There are three 5' exons, one specific to the
CC oocyte (1c), one specific to the pachytene spermatocyte and also
CC found in skeletal muscle (1b) and one found in somatic cells (1a).
CC Three different mRNAs can be produced which give rise to two
CC different translation products: isoform 1 (mRNAs-1a) and isoform 2
CC (mRNA-1b or -1c).
CC
CC -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC
CC -!- SIMILARITY: CONTAINS 1 CXXC-TYPE ZINC FINGER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X14805; CAA32910.1; -
CC EMBL; AF175432; AAF97695.1; -
CC EMBL; AF162282; AAF19352.1; -
CC EMBL; AF175431; AAF60965.1; -
CC EMBL; AF175412; AAF60965.1; JOINED.
CC EMBL; AF175413; AAF60965.1; JOINED.
CC
CC Query Match 97.7%; Score 42; DB 1; Length 1620;
CC Best Local Similarity 88.9%; Pred. No. 2.1;
CC Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1 GLIEKNIEL 9
CC |||||:|
CC Db 431 GLIEKNVEL 439
CC
CC RESULT 3
CC DNML_RAT
CC ID DNML_RAT STANDARD; PRT; 1622 AA.
CC AC Q9ZJ30; Q9WTX3; P70487; Q9WU57; Q9R252;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE DNA (cytosine-5)-methyltransferase 1 (EC 2.1.1.37) (Dnmt1) (DNA
CC methyltransferase I) (DNA MTase RnoIP) (MCMT) (M.RnoIP).
CC GN DNMT1.
CC OS Rattus norvegicus (Rat).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC OX NCBI_TaxID=10116;
CC RN [1]
CC RP SEQUENCE FROM N.A. (ISOFORM 0).
CC RC STRAIN-Sprague-Dawley; TISSUE=Brain, and Placenta;
CC RX MEDLINE=99097263; PubMed=9878564;
CC RA Kimura H., Takeda T., Tanaka S., Ogawa T., Shiota K.;
CC RT "Expression of rat DNA (cytosine-5) methyltransferase (DNA MTase) in
CC rodent trophoblast giant cells: molecular cloning and characterization
CC of rat DNA MTase."
CC RL Biochem. Biophys. Res. Commun. 253:495-501(1998).
CC [2]
CC RP SEQUENCE OF 1-144 FROM N.A. (ISOFORMS 0 AND 8).
CC RC TISSUE=Brain;
CC RA Deng J., Szfy M.;
CC RT "Multiple N-terminal isoforms of DNA (cytosine-5)-methyltransferase
```

```
RT in vivo."
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 17-356 FROM N.A., AND IN VITRO BINDING TO ANNEXIN V.
RC STRAIN=Wistar; TISSUE=Brain;
RX MEDLINE=96301899; PubMed=8667030;
RA Ohsawa K., Imai Y., Ito D., Kohsaka S.;
RT "Molecular cloning and characterization of annexin V-binding proteins
RT with highly hydrophilic peptide structure."
RL J. Neurochem. 67:89-97(1996).
RN [4]
RP SEQUENCE OF 1169-1517 FROM N.A. (ISOFORMS 0; 1; 2; 3; 4; 5; 6 AND 7).
RX MEDLINE=98389705; PubMed=9722504;
RA Deng J., Szfy M.;
RT "Multiple isoforms of DNA methyltransferase are encoded by the
RT vertebrate cytosine DNA methyltransferase gene."
RL J. Biol. Chem. 273:22869-22872(1998).
CC -!- FUNCTION: Methylates CpG residues. Preferentially methylates
CC hemimethylated DNA. It is responsible for maintaining methylation
CC patterns established in development (By similarity). Mediates
CC transcriptional repression by direct binding to HDAC2 (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA -> S-adenosyl-L-
CC homocysteine + DNA containing 5-methylcytosine.
CC -!- SUBUNIT: Binds annexin V (Potential).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: At least 9 isoforms; 0 (shown here), 1/SF1,
CC 2/SF2, 3/SF3, 4/SF4, 5/SF5, 6/SF6, 7/SF7 and 8/short; are produced
CC by alternative splicing
CC -!- TISSUE SPECIFICITY: Isoforms 0 and 8 are highly expressed in
CC placenta, brain, lung, spleen, kidney, heart, and at much lower
CC levels in liver. Isoform 1 is expressed in cerebellum, isoform
CC 2 in muscle and testis, isoform 3 in lung, isoform 4 in spleen and
CC brain, and isoform 5 in brain.
CC -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC -!- SIMILARITY: CONTAINS 2 BAH DOMAINS.
CC -!- SIMILARITY: CONTAINS 1 CXXC-TYPE ZINC FINGER.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB012214; BAA37118.1; -
CC EMBL; AF116344; AAD32541.1; -
CC EMBL; AF116345; AAD32542.1; -
CC EMBL; D64060; BAA20854.1; -
CC EMBL; AF083043; AAD28102.1; -
CC EMBL; AF083038; AAD28102.1; JOINED.
CC EMBL; AF083039; AAD28102.1; JOINED.
CC EMBL; AF083040; AAD28102.1; JOINED.
CC EMBL; AF083041; AAD28102.1; JOINED.
CC EMBL; AF083042; AAD28102.1; JOINED.
CC REBASE; 3019; M.RnoIP.
CC InterPro; IPR001025; BAH.
CC InterPro; IPR001525; C5_DNA_meth.
CC InterPro; IPR002857; Znf_CXXC.
CC Pfam; PF00145; DNA_methylase; 1.
CC Pfam; PF01426; BAH_2.
CC Pfam; PF02008; zf-CXXC; 1.
CC TIGRFAMs; TIGR00675; dcm; 1.
CC SMART; SM00439; BAH; 2.
CC PROSITE; PS00094; C5_MTASE_1; 1.
CC PROSITE; PS00095; C5_MTASE_2; 1.
CC Transferase; Methyltransferase; Transcription regulation; Repressor;
CC DNA-binding; Zinc-finger; Zinc; Metal-binding; Nuclear protein;
CC Repeat; Phosphorylation; Alternative splicing.
CC DOMAIN 173 200 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC FT DOMAIN 327 556 DNA REPLICATION FOCI-TARGETING SEQUENCE.
CC ZN_FING 649 695 CXXC-TYPE.
```


RA Gockel G., Hachtel W.;
 RT "Complete gene map of the plastid genome of the nonphotosynthetic
 RL euglenoid flagellate *Ascia longica*.";
 CC Prolist 151:347-351(2000).
 CC -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: CONTAINS 1 S4 RNA-BINDING DOMAIN.
 CC -!- SIMILARITY: BELONGS TO THE S4P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ294725; CAC24613.1; -.
 DR HSP: P81288; 1C05.
 DR InterPro: IPR001912; Ribosomal_S4.
 DR Pfam: PF00163; Ribosomal_S4; 1.
 DR Pfam: PF01479; S4; 1.
 DR SMART: SM00363; S4; 1.
 DR TIGRFAMs: TIGR01017; rpsd_bact; 1.
 DR PROSITE: PS00632; RIBOSOMAL_S4; 1.
 KW Ribosomal protein; rRNA-binding; Chloroplast.
 FT DOMAIN 93 140 RNA-BINDING (S4 TYPE).
 SQ SEQUENCE 206 AA; 24323 MW; 56FD85EEF8A4B90F CRC64;
 Query Match 81.4%; Score 35; DB 1; Length 206;
 Best Local Similarity 87.5%; Pred. No. 6.6;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 2 LIEKNIEL 9
 DB 151 IIEKNIEL 158
 RESULT 6
 ID PU92_YEAST STANDARD; PRT; 592 AA.
 AC P38009;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Bifunctional purine biosynthesis protein ADEL7 [Includes:
 DE phosphoribosylaminoimidazolecarboxamide formyltransferase (EC 2.1.2.3)
 DE (AICAR transferase); IMP cyclohydrolase (EC 3.5.4.10) (inosinidase)
 DE (IMP synthetase) (ATIC)].
 GN ADEL7 OR YMR120C OR YMR564.02C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 RN NCBI_TaxID=4932;
 RP SEQUENCE FROM N.A.
 RA Podivinsky E., Barrell B.G., Rajandream M.A.;
 RC STRAIN=S288C / AB972;
 CC Lye G., Churcher C.M., Barrell B.G., Rajandream M.A.;
 RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE OF 140-150 AND 389-400.
 RC STRAIN=S288C;
 RX MEDLINE=95203288; PubMed=7895733;
 RA Garrels J.I., Futcher B., Kobayashi R., Latter G.I., Schwender B.,
 RA Volpe T., Warner J.R., McLaughlin C.S.;
 RT "Protein identifications for a Saccharomyces cerevisiae protein
 RT database.";
 RL Electrophoresis 15:1466-1486(1994).
 CC -!- CATALYTIC ACTIVITY: 10-formyltetrahydrofolate + 5-amino-1-(5-
 CC phospho-D-ribose)imidazole-4-carboxamide = tetrahydrofolate + 5-
 CC formamido-1-(5-phospho-D-ribose)imidazole-4-carboxamide.
 CC -!- CATALYTIC ACTIVITY: IMP + H(2)O = 5-formamido-1-(5-phospho-D-

CC ribosyl)imidazole-4-carboxamide.
 CC -!- PATHWAY: De novo purine biosynthesis; ninth step.
 CC -!- PATHWAY: De novo purine biosynthesis; tenth step.
 CC -!- SUBUNIT: HOMODIMER (POSSIBLE).
 CC -!- DOMAIN: THE IMP CYCLOHYDROLASE ACTIVITY RESIDES IN THE N-TERMINAL
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE PURH FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Z49273; CAA89269.1; -.
 DR SGD: S0004727; ADEL7.
 DR InterPro: IPR002695; AICARFT_IMPCHas.
 DR InterPro: IPR004362; MGS_Like.
 DR Pfam: PF01808; AICARFT_IMPCHas; 1.
 DR Pfam: PF02142; MGS; 1.
 DR ProDom: PD004666; AICARFT_IMPCHas; 1.
 DR TIGRFAMs: TIGR00355; purH; 1.
 DR Purine biosynthesis; transferase; Hydrolase; Multifunctional enzyme.
 KW Purine biosynthesis; transferase; Hydrolase; Multifunctional enzyme.
 FT CONFLICT 389 389 R -> A (IN REF. 2).
 SQ SEQUENCE 592 AA; 65263 MW; 8ABA71761B512242 CRC64;
 Query Match 81.4%; Score 35; DB 1; Length 592;
 Best Local Similarity 66.7%; Pred. No. 19;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GLIEKNIEL 9
 DB 24 GLIEKNIEL 32
 RESULT 7
 ID CHSY_MALDO STANDARD; PRT; 232 AA.
 AC P30078;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase (EC 2.3.1.74) (Naregenin-chalcone synthase)
 DE (Fragment).
 GN CHS.
 OS Malus domestica (Apple) (Malus sylvestris).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
 CC NCBI_TaxID=3750;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Leaf.
 RA Podivinsky E., Bradley J.M., Davis K.M.;
 RL Submitted (AUG-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN
 CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X68977; CAA48773.1; -
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; PARTIAL.
KW Flavonoid biosynthesis; Transferase; Acyltransferase.
FT NON_TER 1
FT ACT_SITE 7
SQ SEQUENCE 232 AA; 24616 MW; 12D6113C80A9E86B CRC64;

Query Match 79.1%; Score 34; DB 1; Length 232;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GLIEKNIE 8
Db 116 GLISKNIE 123
||| ||||
BY SIMILARITY.

RESULT 8
CHS1_SOYBN
ID CHS1_SOYBN STANDARD; PRT; 388 AA.
AC P24826;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).
GN CHS1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RX MEDLINE=91329712; PubMed=1868209;
RA Akada S.; Kung S.D.; Dube S.K.;
RT "The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene family.";
RL Plant Mol. Biol. 16:751-752(1991).
CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.
CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X54644; CAA38456.1; -
DR PIR; S15006; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X68977; CAA48773.1; -
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; PARTIAL.
KW Flavonoid biosynthesis; Transferase; Acyltransferase.
FT NON_TER 1
FT ACT_SITE 7
SQ SEQUENCE 232 AA; 24616 MW; 12D6113C80A9E86B CRC64;

Query Match 79.1%; Score 34; DB 1; Length 232;
Best Local Similarity 87.5%; Pred. No. 12;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GLIEKNIE 8
Db 116 GLISKNIE 123
||| ||||
BY SIMILARITY.

RESULT 8
CHS1_SOYBN
ID CHS1_SOYBN STANDARD; PRT; 388 AA.
AC P24826;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).
GN CHS1.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RX MEDLINE=91329712; PubMed=1868209;
RA Akada S.; Kung S.D.; Dube S.K.;
RT "The nucleotide sequence of gene 1 of the soybean chalcone synthase multigene family.";
RL Plant Mol. Biol. 16:751-752(1991).
CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.
CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X54644; CAA38456.1; -
DR PIR; S15006; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X68977; CAA48773.1; -
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;

KW Multigene family.
FT ACT_SITE 164
SQ SEQUENCE 388 AA; 42516 MW; 73AC3B59A4E91BE1 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 388;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
Db 272 GLISKNIE 279
||| ||||
BY SIMILARITY.

RESULT 9
CHS2_SOYBN
ID CHS2_SOYBN STANDARD; PRT; 388 AA.
AC P17957;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
GN CHS2.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RX MEDLINE=90287722; PubMed=2356130;
RA Akada S.; Kung S.D.; Dube S.K.;
RT "Nucleotide sequence of one member of soybean chalcone synthase multi-gene family.";
RL Nucleic Acids Res. 18:3398-3398(1990).
CC [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Williams;
RX Akada S.; Kung S.D.; Dube S.K.;
RL Submitted (APR 1992) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE) WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO NARINGENIN.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA + naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY OF WHICH ARE BRIGHTLY COLORED.
CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL: X52097; CAA36317.1; -
DR EMBL: X65636; CAA46590.1; -
DR PIR; S10475; SYSYCN.
DR InterPro: IPR001099; N-C synthase.
DR Pfam: PF00195; Chal_stil_syntC; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Multigene family.
FT ACT_SITE 164
FT CONFLICT 9
FT CONFLICT 100
FT CONFLICT 299
FT CONFLICT 299

FT CONFLICT 387 387 T -> S (IN REF. 2).
SQ SEQUENCE 388 AA; 42504 MW; 634D3F1CDC5F973 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 388;

Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIE 8
||| ||||
Db 272 GLISKNIIE 279

RESULT 10

CHS3_SOYBN STANDARD; PRT; 388 AA.
AC P19168;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 3 (EC 2.3.1.74) (Naringenin-chalcone synthase 3).
GN CHS3.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Williams;
RX MEDLINE=91016949; PubMed=2216793;
RA Akada S., Kung S.D., Dube S.K.;
RT "The nucleotide sequence of gene 3 of the soybean chalcone synthase
RT multigene family.";
RL Nucleic Acids Res. 18:5899-5899(1990).
CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X53958; CAA37909.1; -.
PIR; S11486; SYSYC3.
DR InterPro; IPR001099; N-C_synthase.
DR Pfam; PF00195; Chal_stil_syntc; 1.
DR Pfam; PF02797; Chal_stil_syntc; 1.
DR ProDom; PD000453; N-C_synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;
KW Multigene family.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 388 AA; 42389 MW; 00FFD0982F2B6B0 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 388;

Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIE 8
||| ||||
Db 272 GLISKNIIE 279

RESULT 11

CHS5_SOYBN STANDARD; PRT; 388 AA.
AC P48406;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 5 (EC 2.3.1.74) (Naringenin-chalcone synthase 5).
GN CHS5.
OS Glycine max (Soybean).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
OX NCBI_TaxID=3847;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Williams;
RX MEDLINE=96046740; PubMed=7579172;
RA Akada S., Dube S.K.;
RT "Organization of soybean chalcone synthase gene clusters and
RT characterization of a new member of the family.";
RL Plant Mol. Biol. 29:189-199(1995).
CC -!- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
CC NARINGENIN.
CC -!- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
CC naringenin-chalcone + 3 CO(2).
CC -!- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
CC OF WHICH ARE BRIGHTLY COLORED.
CC -!- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; L07647; AAB01004.1; -.
DR InterPro; IPR001099; N-C_synthase.
DR Pfam; PF00195; Chal_stil_syntc; 1.
DR Pfam; PF02797; Chal_stil_syntc; 1.
DR ProDom; PD000453; N-C_synthase; 1.
DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;
KW Multigene family.
FT ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 388 AA; 42534 MW; F924427A75FF1C20 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 388;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIE 8
||| ||||
Db 272 GLISKNIIE 279

RESULT 12

CHS1_CAMSI STANDARD; PRT; 389 AA.
AC P48386;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chalcone synthase 1 (EC 2.3.1.74) (Naringenin-chalcone synthase 1).
GN CHS1.
OS Camellia sinensis (Tea).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; Ericales; Theaceae; Camellia.
NCBI_TaxID=4442;
[1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Yabukita; TISSUE=Leaf;
RX MEDLINE=95120283; PubMed=7820373;
RA Takeuchi A., Matsumoto S., Hayatsu M.;
RT "Chalcone synthase from Camellia sinensis: isolation of the cDNAs and
the organ-specific and sugar-responsive expression of the genes.";
RL Plant Cell Physiol. 35;1011-1018(1994).
CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
NARINGENIN.
CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
naringenin-chalcone + 3 CO(2).
CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
FLAVONOIDS. A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
OF WHICH ARE BRIGHTLY COLORED.
CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL: D26593; BAA05640.1; -
DR InterPro: IPR001099; N-C_synthase.
DR Pfam: PF00195; Chal_stil_synt; 1.
DR Pfam: PF02797; Chal_stil_syntC; 1.
DR ProDom: PD000453; N-C_synthase; 1.
DR PROSITE: PS00441; CHALCONE_SYNTH; 1.
KW Flavonoid biosynthesis; Transferase; Acyltransferase;
FT Multigene family. 164
ACT_SITE 164 164 BY SIMILARITY.
SQ SEQUENCE 389 AA; 42570 MW; DCAA508258C3F973 CRC64;

Query Match 79.1%; Score 34; DB 1; Length 389;
Best Local Similarity 87.5%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
DB 273 GLISKNIE 280
||| ||||

RESULT 13
CHSL_CASGL
ID CHSL_CASGL STANDARD; PRT; 389 AA.
AC Q9ZRR8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chalcone synthase (EC 2.3.1.74) (Naringenin-chalcone synthase).
GN CHSL.
OS Casuarina glauca (Swamp oak).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Casuarinaceae; Casuarina.
NCBI_TaxID=3522;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Root nodules;
RX PubMed=10492666;
RA Laplace L., Gherbi H., Frutz T., Pawlowski K., Franche C.,
RA Macheix J.J., Auguy F., Bogusz D., Duhoux E.;
RT "Flavan-containing cells delimit Frankia infected compartments in
Casuarina glauca nodules.";
RL Plant Physiol. 121:113-122(1999).

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; X55194; CAA38980.1; -.
 DR PIR; S12223.
 DR InterPro: IPR001099; N-C-synthase.
 DR Pfam; PF00195; Chal_stil_syntC; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR ProDom; PD000453; N-C-synthase; 1.
 DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase;
 KW Multigene family.
 FT ACT_SITE 164 BY SIMILARITY.
 SQ SEQUENCE 389 AA; 42552 MW; 553DC69E5EA96A8B CRC64;

Query Match 79.1%; Score 34; DB 1; Length 389;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 GLIEKNIE 8
 ||| ||||
 Db 273 GLISKNI 280

RESULT 15

CHS2_CAMSI STANDARD; PRT; 389 AA.
 ID CHS2_CAMSI
 AC P48387;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Chalcone synthase 2 (EC 2.3.1.74) (Naringenin-chalcone synthase 2).
 GN CHS2.
 OS Camellia sinensis (Tea).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; Ericales; Theaceae; Camellia.
 OX NCBI_TaxID=4442;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Yabukita; TISSUE=Leaf;
 RX MEDLINE=95120283; PubMed=7820373;
 RA Takeuchi A., Matsumoto S., Hayatsu M.;
 RT "Chalcone synthase from Camellia sinensis: Isolation of the cDNAs and
 RT the organ-specific and sugar-responsive expression of the genes.";
 RL Plant Cell Physiol. 35:1011-1018(1994).
 CC -1- FUNCTION: THE PRIMARY PRODUCT OF THIS ENZYME IS 4,2',4',6'-
 CC TETRAHYDROXYCHALCONE (ALSO TERMED NARINGENIN-CHALCONE OR CHALCONE)
 CC WHICH CAN UNDER SPECIFIC CONDITIONS SPONTANEOUSLY ISOMERIZE INTO
 CC NARINGENIN.
 CC -1- CATALYTIC ACTIVITY: 3 malonyl-CoA + 4-coumaroyl-CoA = 4 CoA +
 CC naringenin-chalcone + 3 CO(2).
 CC -1- PATHWAY: PART OF THE BIOSYNTHETIC PATHWAY FOR ALL CLASSES OF
 CC FLAVONOIDS, A LARGE CLASS OF SECONDARY PLANT METABOLITES, MANY
 CC OF WHICH ARE BRIGHTLY COLORED.
 CC -1- SIMILARITY: BELONGS TO THE CHALCONE/STILBENE SYNTHASES FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; D26594; BAA05641.1; -.
 DR InterPro: IPR001099; N-C-synthase.
 DR Pfam; PF00195; Chal_stil_syntC; 1.
 DR Pfam; PF02797; Chal_stil_syntC; 1.
 DR ProDom; PD000453; N-C-synthase; 1.

DR PROSITE; PS00441; CHALCONE_SYNTH; 1.
 KW Flavonoid biosynthesis; Transferase; Acyltransferase;
 KW Multigene family.
 FT ACT_SITE 164 BY SIMILARITY.
 SQ SEQUENCE 389 AA; 42595 MW; 74ACC577956F9DBA CRC64;
 Query Match 79.1%; Score 34; DB 1; Length 389;
 Best Local Similarity 87.5%; Pred. No. 20;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 GLIEKNIE 8
 ||| ||||
 Db 273 GLISKNI 280

Search completed: February 4, 2003, 17:36:55
 Job time : 11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2003, 17:36:35 ; Search time 28 Seconds
(without alignments)
66.229 Million cell updates/sec

Title: US-09-865-548A-13
Perfect score: 43
Sequence: 1 GLIEKNIEL 9

Scoring Table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42	97.7	1490	13 P79922	P79922 xenopus lae
2	42	97.7	1499	13 Q8QCB8	Q8qcb8 brachydanio
3	42	97.7	1501	10 Q9PR60	Q9pr60 oryza sativ
4	42	97.7	1503	13 Q918X6	Q918x6 xiphophorus
5	37	86.0	152	17 Q972N7	Q972n7 sulfolobus
6	35	81.4	265	17 Q97WN2	Q97wn2 sulfolobus
7	35	81.4	269	5 Q9VKV4	Q9vkv4 drosophila
8	35	81.4	350	17 Q8U3Y5	Q8u3y5 pyrococcus
9	35	81.4	395	13 Q9PWG6	Q9pwg6 anguilla ja
10	35	81.4	408	16 Q9RGT7	Q9rgt7 mycoplasma
11	35	81.4	425	17 Q8TWY9	Q8twy9 methanopyru
12	34	79.1	147	10 Q8RVM6	Q8rvm6 malus domes
13	34	79.1	147	10 Q8RVM5	Q8rvm5 malus domes
14	34	79.1	184	10 Q9LEH0	Q9leh0 juglans nig
15	34	79.1	184	10 Q94JN8	Q94jn8 cucumis sat
16	34	79.1	190	10 Q93XC7	Q93xc7 olea europa

17 34 79.1 253 10 Q9LKN5
18 34 79.1 277 10 Q948T7
19 34 79.1 281 10 Q8S993
20 34 79.1 318 10 Q23885
21 34 79.1 323 17 Q58571
22 34 79.1 333 10 Q9AVC1
23 34 79.1 333 10 Q9AVC0
24 34 79.1 374 10 Q81476
25 34 79.1 379 10 Q04065
26 34 79.1 386 10 Q23923
27 34 79.1 389 10 Q9FEY5
28 34 79.1 389 10 Q9M5B2
29 34 79.1 389 10 Q93XP8
30 34 79.1 389 10 Q42865
31 34 79.1 389 10 Q42864
32 34 79.1 389 10 Q43040
33 34 79.1 389 10 Q8W3P6
34 34 79.1 389 10 Q8S4Y8
35 34 79.1 389 10 Q8S4Y6
36 34 79.1 389 10 Q8RVK9
37 34 79.1 390 10 Q48564
38 34 79.1 390 10 Q22122
39 34 79.1 390 10 Q8S4Y7
40 34 79.1 391 10 Q9SEP4
41 34 79.1 391 10 Q9SEP3
42 34 79.1 391 10 Q9AU11
43 34 79.1 391 10 Q9AU10
44 34 79.1 391 10 Q9AU09
45 34 79.1 392 10 Q9SEP5

ALIGNMENTS

RESULT 1

P79922 ID P79922 PRELIMINARY; PRT; 1490 AA.
AC P79922;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Modification methylase (EC 2.1.1.73) (Cytosine-specific methyltransferase).
DE DNA METASE.
GN Xenopus laevis (African clawed frog).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=OVARY;
RX MEDLINE=97164021; PubMed=9010768;
RA Kimura H., Ishihara G., Tajima S.;
RT "Isolation and expression of a Xenopus laevis DNA Methyltransferase cDNA".
RL J. Biochem. 120:1182-1189(1996).
RL -!- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.
CC -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
DR EMBL; D78638; BA011458.1;
DR REBASE; 2998; M.XIai.
DR InterPro; IPR001025; BAH.
DR InterPro; IPR001525; C5_DNA_meth.
DR InterPro; IPR002857; znf_CXXC.
DR Pfam; PF01426; BAH; 2.
DR Pfam; PF00145; DNA_methylase; 1.
DR Pfam; PF02008; zf-CXXC; 1.
DR PRINTS; PR00105; C5METTRFRASE.
DR SMART; SM00439; BAH; 2.
DR TIGRFAMS; TIGR00675; dcm; 1.
DR PROSITE; PS00094; C5_MTASE_1; 1.
DR PROSITE; PS00095; C5_MTASE_2; 1.

```

KW Methyltransferase; Transferase.
SQ SEQUENCE 1490 AA; 167983 MW; 1AF48E0CC11D906D CRC64;

Query Match          97.7%; Score 42; DB 13; Length 1490;
Best Local Similarity 88.9%; Pred. No. 9.2;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
   |||||:|
DB 300 GLIEKNVEL 308

RESULT 2
Q9QGB8
ID Q9QGB8 PRELIMINARY; PRT; 1499 AA.
AC Q9QGB8;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE DNA (Cytosine-5)-methyltransferase.
GN DMNT1.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21426831; PubMed=11536427;
RA Mhanni A., Yoder J., Dubesky C., McGowan R.A.;
RT "Cloning and sequence analysis of a zebrafish cDNA encoding DNA
   (cytosine-5)-methyltransferase-1.";
RL Genesis 30:213-219(2001).
RN [2]
RP SEQUENCE FROM N.A.
RA Mhanni A., Yoder J., Dubesky C., McGowan R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF483203; AAL86596.1; -.
KW Transferase; Methyltransferase.
SQ SEQUENCE 1499 AA; 168803 MW; 8AA457B45DDE29AB CRC64;

Query Match          97.7%; Score 42; DB 13; Length 1499;
Best Local Similarity 88.9%; Pred. No. 9.3;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
   |||||:|
DB 306 GLIEKNVEL 314

RESULT 3
Q9FR60
ID Q9FR60 PRELIMINARY; PRT; 1501 AA.
AC Q9FR60;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Modification methylase (EC 2.1.1.73) (Cytosine-specific
   methyltransferase).
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. INDICA;
RA Wiriyakorn N., Narangajavana J.;
RT "Isolation and characterization of the putative cDNA encoding cytosine
   DNA methyltransferase from rice (Oryza sativa).";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-
   ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.

```

```

CC -|- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
DR EMBL; AF155874; BAG43376.1; -.
DR InterPro; IPR001025; BAH.
DR InterPro; IPR001525; C5_DNA_meth.
DR InterPro; IPR002857; znf_CXXC.
DR Pfam; PF01426; BAH; 2.
DR Pfam; PF00145; DNA_methylase; 1.
DR Pfam; PF02008; zf-CXXC; 1.
DR PRINTS; PR00105; C5METTRFRASE.
DR SMART; SM00439; BAH; 2.
DR TIGRFAMS; TIGR00675; dcm; 1.
DR PROSITE; PS00094; C5_MTASE_1; 1.
DR PROSITE; PS00095; C5_MTASE_2; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 1501 AA; 169588 MW; BDCF5D1C410BA3BD CRC64;

Query Match          97.7%; Score 42; DB 10; Length 1501;
Best Local Similarity 88.9%; Pred. No. 9.3;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
   |||||:|
DB 312 GLIEKNVEL 320

RESULT 4
Q918X6
ID Q918X6 PRELIMINARY; PRT; 1503 AA.
AC Q918X6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Modification methylase (EC 2.1.1.73) (Cytosine-specific
   methyltransferase).
GN DNMT-1.
OS Xiphophorus maculatus x Xiphophorus helleri.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Cyprinodontiformes; Poeciliidae; Xiphophorus.
OX NCBI_TaxID=96538;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20293049; PubMed=10831840;
RA Altschmid J., Volff J.N., Winkler C., Gutbrod H., Korting C.,
RA Pagany M., Schartl M.;
RT "Primary structure and expression of the Xiphophorus DNA-(cytosine-5)-
   methyltransferase XDNMT-1.";
RL Gene 249:75-82(2000).
CC -|- CATALYTIC ACTIVITY: S-ADENOSYL-L-METHIONINE + DNA CYTOSINE = S-
   ADENOSYL-L-HOMOCYSTEINE + DNA 5-METHYLCYTOSINE.
CC -|- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
DR EMBL; AF152342; AAF73200.1; -.
DR InterPro; IPR001025; BAH.
DR InterPro; IPR001525; C5_DNA_meth.
DR InterPro; IPR002857; znf_CXXC.
DR Pfam; PF01426; BAH; 2.
DR Pfam; PF00145; DNA_methylase; 1.
DR Pfam; PF02008; zf-CXXC; 1.
DR PRINTS; PR00105; C5METTRFRASE.
DR SMART; SM00439; BAH; 2.
DR TIGRFAMS; TIGR00675; dcm; 1.
DR PROSITE; PS00094; C5_MTASE_1; 1.
DR PROSITE; PS00095; C5_MTASE_2; 1.
KW Methyltransferase; Transferase.
SQ SEQUENCE 1503 AA; 169575 MW; E5E50C3AECE0E330 CRC64;

Query Match          97.7%; Score 42; DB 13; Length 1503;
Best Local Similarity 88.9%; Pred. No. 9.3;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9
   |||||:|

```

Db 307 GLIEKNVEL 315

RESULT 5

Q972N7

ID Q972N7 PRELIMINARY; PRT; 152 AA.

AC Q972N7

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Hypothetical protein ST1095.

GN ST1095.

OS Sulfolobus tokodaii.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=111955;

[1]

SEQUENCE FROM N.A.

RP STRAIN=JCM 10545 / 7;

RC PubMed=11572479;

RX Kawabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y., Yoshizawa T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Masuda S., Yanagih M., Nishimura M., Yamagishi A., Oshima T., Kikuchi H.;

RA "Complete genome sequence of an aerobic thermoacidophilic Crenarchaeon, Sulfolobus tokodaii strain 7.";

RT DNA Res. 8:123-140(2001).

RL EMBL; AP000985; BAB66127.1;

DR InterPro; IPR000572; Euk Mb.oxred.

DR Pfam; PF00174; oxidored_molyb; 1.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 152 AA; 17514 MW; 34DF62C2DDC7F5AE CRC64;

Query Match 86.0%; Score 37; DB 17; Length 152;

Best Local Similarity 77.8%; Pred. No. 10;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIEL 9

Db 10 GLVEKNINL 18

RESULT 6

Q97WN2

ID Q97WN2 PRELIMINARY; PRT; 265 AA.

AC Q97WN2

DT 01-OCT-2001 (TREMBLrel. 18, Created)

DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE Hypothetical protein SSO2179.

GN SSO2179.

OS Sulfolobus solfataricus.

OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;

OC Sulfolobus.

OX NCBI_TaxID=2287;

[1]

SEQUENCE FROM N.A.

RP STRAIN=ATCC 35092 / DSM 1617 / P2;

RC MEDLINE=21332296; PubMed=11427726;

RX She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G., Aweye M.J., Chan-Weiher C.C.-Y., Clausen I.G., Curtis B.A., De Moors A., Erauso G., Fletcher C., Gordon P.M.K., Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X., Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N., Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T., Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;

RA "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";

RT Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

RL EMBL; AE006824; AAK42354.1;

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 265 AA; 31527 MW; F8D949DA90F5B362 CRC64;

Query Match 81.4%; Score 35; DB 17; Length 265;

Best Local Similarity 75.0%; Pred. No. 46;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8

Db 93 GILEKNIE 100

RESULT 7

Q9VKV4

ID Q9VKV4 PRELIMINARY; PRT; 269 AA.

AC Q9VKV4

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE CGI13142 protein.

GN CGI13142.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephyroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

[1]

SEQUENCE FROM N.A.

RP STRAIN=BERKELEY;

RC MEDLINE=20196006; PubMed=10731132;

RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P., Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.-H., Dew I., Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RA "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195(2000).

DR EMBL; AE003628; AAF52955.1;

DR FlyBase; FBgn0032251; CGI13142.

SQ SEQUENCE 269 AA; 30694 MW; DCAD330F1C2A17C CRC64;

Query Match 81.4%; Score 35; DB 5; Length 269;

Best Local Similarity 87.5%; Pred. No. 47;

Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LIEKNIEL 9
 |||||
 Db 23 LIEKNIEL 30

RESULT 8
 Q8U3Y5
 ID Q8U3Y5 PRELIMINARY; PRT; 350 AA.
 AC Q8U3Y5
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein PF0316.
 GN PF0316
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VCL / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010156; AAL80440.1; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 350 AA; 40187 MW; 4331411F2FBA5341 CRC64;

Query Match 81.4%; Score 35; DB 17; Length 350;
 Best Local Similarity 77.8%; Pred. No. 61;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIEL 9
 |||||
 Db 191 GLIEKNIEL 199

RESULT 9
 Q9PWG6
 ID Q9PWG6 PRELIMINARY; PRT; 395 AA.
 AC Q9PWG6
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Activin B.
 OS Anguilla japonica (Japanese eel).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Anguilliformes; Anguillidae;
 OC Anguilla.
 OX NCBI_TaxID=7937;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-TESTIS;
 RA Miura T., Miura C., Eto Y., Nagahama Y.;
 RT "Activin B gene is required for the initiation of spermatogenesis in the Japanese eel, Anguilla japonica.";
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 DR EMBL; AB025356; BAA83804.1; -
 DR HSSP; P12643; 3BMP.
 DR InterPro; IPR000381; Inhibin_betaB.
 DR InterPro; IPR001318; Inhibin_betaC.
 DR InterPro; IPR001839; TGFb.
 DR InterPro; IPR001111; TGFb_N.
 DR Pfam; PF00019; TGF-beta; 1.
 DR Pfam; PF00688; TGFb_propeptide; 1.
 DR PRINTS; PR00671; INHIBINB.
 DR PRINTS; PR00672; INHIBINB.
 DR ProDom; PD000357; TGFb; 1.
 DR SMART; SM00204; TGFb; 1.
 DR PROSITE; PS00250; TGF_BETA_1; 1.
 KW Glycoprotein.

SQ SEQUENCE 395 AA; 43889 MW; FA56DD62D18509A3 CRC64;

Query Match 81.4%; Score 35; DB 13; Length 395;
 Best Local Similarity 66.7%; Pred. No. 68;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GLIEKNIEL 9
 |||||
 Db 201 GLVEKRVEL 209

RESULT 10
 Q98RG7
 ID Q98RG7 PRELIMINARY; PRT; 408 AA.
 AC Q98RG7
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
 DE Hypothetical protein MYPU_0420.
 GN MYPU_0420.
 OS Mycoplasma pulmonis.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
 OC Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=2107;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-UAB CTIP;
 RX MEDLINE=21267165; PubMed=11353084;
 RA Chamberaud I., Heilig R., Ferris S., Barbe V., Samson D., Gallissou F., Mosser I., Dybdvig K., Wroblewski H., Viari A., Rocha E.P.C., Blanchard A.;
 RT "The complete genome sequence of the murine respiratory pathogen Mycoplasma pulmonis.";
 RL Nucleic Acids Res. 29:2145-2153(2001).
 DR EMBL; AL445563; CAC13215.1; -
 DR Mypulist; MYPU_0420; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 408 AA; 49326 MW; 1C3566880A65ACE0 CRC64;

Query Match 81.4%; Score 35; DB 16; Length 408;
 Best Local Similarity 87.5%; Pred. No. 70;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 LIEKNIEL 9
 |||||
 Db 230 LIEKNIEL 237

RESULT 11
 Q8TWY9
 ID Q8TWY9 PRELIMINARY; PRT; 425 AA.
 AC Q8TWY9
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Translation initiation factor eIF2B subunit.
 GN MK0892.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE=21927647; PubMed=11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N., Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L., Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O., Malykh A.G., Koonin E.V., Kozyavkin S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19 and monophyly of archaeal methanogens.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL; AE010378; RAN02105.1; -

```
KW Initiation factor; Complete proteome.
SQ SEQUENCE 425 AA; 46543 MW; 2C21B47B7E97B870 CRC64;

Query Match      81.4%; Score 35; DB 17; Length 425;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GLIEKNIE 9
   ||| |||
Db 245 GVIEENVEL 253

RESULT 12
Q8RVM6 PRELIMINARY; PRT; 147 AA.
AC Q8RVM6;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Chalcone synthase (Fragment).
GN CHS-1.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EVERESTE X MM106; TISSUE=LEAF;
RA Venisse J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF494401; AAM12894.1; -
DR NON_TER 1
FT NON_TER 147
SQ SEQUENCE 147 AA; 15564 MW; 2A0DF7A701F118B CRC64;

Query Match      79.1%; Score 34; DB 10; Length 147;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
   ||| |||
Db 81 GLISKNIIE 88

RESULT 13
Q8RVM5 PRELIMINARY; PRT; 147 AA.
AC Q8RVM5;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE Chalcone synthase (Fragment).
GN CHS-2.
OS Malus domestica (Apple) (Malus sylvestris).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX NCBI_TaxID=3750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. EVERESTE X MM106; TISSUE=LEAF;
RA Venisse J.-S., Malnoy M., Faize M., Paulin J.-P., Brisset M.-N.;
RT "Modulation of defense responses of Malus during incompatible and
compatible interactions with Erwinia amylovora.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF494402; AAM12895.1; -
DR NON_TER 1
FT NON_TER 147
SQ SEQUENCE 147 AA; 15656 MW; 38082000518D0E95 CRC64;
```

```
Query Match      79.1%; Score 34; DB 10; Length 147;
Best Local Similarity 87.5%; Pred. No. 41;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
   ||| |||
Db 81 GLISKNIIE 88

RESULT 14
Q9LEH0 PRELIMINARY; PRT; 184 AA.
AC Q9LEH0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE Chalcone synthase (EC 2.3.1.74) (Fragment).
GN CHS.
OS Juglans nigra (black walnut).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Juglandaceae; Juglans.
OX NCBI_TaxID=16719;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=DIFFERENTIATING SECONDARY XYLEM;
RA Beritognolo I., Breton C., Jay-Allemand C.;
RT "Cloning of cDNA sequences encoding structural genes of
phenylpropanoid pathway in Juglans nigra L.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278456; CAB97495.1; -
DR InterPro; IPR001099; N-C-synthase.
DR Pfam; PF00195; Chal_stil_synth; 1.
DR Pfam; PF02797; Chal_stil_synthC; 1.
DR ProDom; PD000453; N-C-synthase; 1.
DR KW Acyltransferase; Transferase.
FT NON_TER 1
FT NON_TER 184
SQ SEQUENCE 184 AA; 19674 MW; 599DCA4927FB6A16 CRC64;

Query Match      79.1%; Score 34; DB 10; Length 184;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GLIEKNIE 8
   ||| |||
Db 116 GLISKNIIE 123

RESULT 15
Q94JN8 PRELIMINARY; PRT; 184 AA.
AC Q94JN8;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE Chalcone synthase (Fragment).
OS Cucumis sativus (Cucumber).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.
OX NCBI_TaxID=3659;
RN [1]
RP SEQUENCE FROM N.A.
RA Forana B., Morency M.J., Belanger R.R., Seguin A.;
RT "Elicitor induced resistance in powdery mildew infected cucumber.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF376133; AAK54648.1; -
DR InterPro; IPR001099; N-C-synthase.
DR Pfam; PF00195; Chal_stil_synth; 1.
DR Pfam; PF02797; Chal_stil_synthC; 1.
DR ProDom; PD000453; N-C-synthase; 1.
FT NON_TER 1
```

FT NON_TER 184 184
SQ SEQUENCE 184 AA; 19760 MW; 02A6E60E0862BEF7 CRC64;
Query Match 79.1%; Score 34; DB 10; Length 184;
Best Local Similarity 87.5%; Pred. No. 52;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GLIEKNIE 8
||| ||||
Db 116 GLISKNIE 123

Search completed: February 4, 2003, 17:37:32
Job time : 30 secs